

Western  Graduate&PostdoctoralStudies

Western University  
**Scholarship@Western**

---

Electronic Thesis and Dissertation Repository

---

8-23-2017 1:00 PM

## Expression and Characterization of a Putative **B**eta-Glycosidase from the Ginseng Pathogen *Pythium Irregulare*

Scarlett Puebla Barragan  
*The University of Western Ontario*

Supervisor  
Mark Bernards  
*The University of Western Ontario*

Graduate Program in Biology

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science  
© Scarlett Puebla Barragan 2017

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

 Part of the [Biology Commons](#)

---

### Recommended Citation

Puebla Barragan, Scarlett, "Expression and Characterization of a Putative Beta-Glycosidase from the Ginseng Pathogen *Pythium Irregulare*" (2017). *Electronic Thesis and Dissertation Repository*. 4799.  
<https://ir.lib.uwo.ca/etd/4799>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

## ABSTRACT

A putative ginsenosidase (PiGH1-x) from the oomycete pathogen of American Ginseng, *Pythium irregulare*, was expressed in *E. coli* DH5 $\alpha$  to assess its capability to partially deglycosylate ginsenosides. The recombinant protein was extracted using non-denaturing and denaturing conditions and purified using gel filtration chromatography and nickel affinity purification. Glycosidase activity was tested using *p*-nitrophenyl  $\beta$ -D-glucopyranoside (pNPG) as the substrate and measured through spectrophotometry. Ginsenosidase activity was tested using ginsenoside Rb1 as the substrate, reaction products were identified using LC-MS. Recombinant PiGH1-x cleaved glucose from pNPG, and converted ginsenoside Rb1 into Rd and Gypenoside XVII, demonstrating both  $\beta(1\rightarrow2)$  and  $\beta(1\rightarrow6)$  glycosidase activity. While a high yield of recombinant protein was observed, protein aggregation interfered with the purification process, and further optimization is required. Since previous work has correlated ginsenoside deglycosylation with infection severity, these findings suggest that PiGH1-x may have an important role in the pathogenicity of *P. irregulare* towards American Ginseng.

**Keywords:** *Pythium irregulare*, American Ginseng, ginsenosides, ginsenosidase, glycosidase.

## DEDICATION

Dedico esta tesis a mis padres, cuyo amor y apoyo incondicional me ha permitido convertirme en la persona y profesionista que soy:

Ana Bertha Barragán Birrueta y José Eugenio Puebla Calderón

## ACKNOWLEDGEMENTS

I would not have been able to complete my MSc without the support I received throughout these two years, therefore I would like to express my gratitude to the following individuals.

First and foremost, to my supervisor, Dr. Mark A. Bernards, for giving me the chance to form part of his amazing work group and for always being there to give his support and encouragement. Thank you for trusting me, and for sharing your knowledge with me. I will be forever grateful for the opportunity you gave me.

I also want to thank the members of my advisory committee, Dr. R. Greg Thorn and Dr. Susanne Kohalmi, for their time and support which made my project more complete.

Dr. John McCormick and to the members of the McCormick lab, with an especial mention to Dr. Lorea Baroja for sharing her knowledge as well as for her endless patience and technical support.

The former members of the Bernards lab, Dimitre Ivanov and Anica Bjelica for sharing their expertise with me and for their assistance.

My labmates during these two years, Katie Woolfson, Brendan Walshe-Roussel and Alejandra Rey for their continuous help and friendship. I am especially grateful to Trish Tully for her guidance and support (both moral and technical) during these past two years.

The members of the Hüner and Trick labs, thank you for your technical support and friendship.

Dr. Silverio García Lara, for his support when I decided to start with graduate school.

My parents for their endless love, and for always supporting me in every possible way no matter the distance, nothing would have been possible without them.

To Sergio Domínguez, not only for sharing with me his technical knowledge, but for being my team-mate in this adventure of getting outside of our comfort zone and coming to the cold and beautiful North.

Finally, I thank and acknowledge the funding I received during my studies from CONACyT, NSERC and Western University.

# TABLE OF CONTENTS

Abstract .....	i
Dedication .....	ii
Acknowledgements .....	iii
Table of contents .....	iv
List of Tables .....	vii
List of Figures .....	viii
List of Abbreviations .....	ix
List of Appendices .....	x
1 Introduction .....	1
1.1 Ginseng overview .....	1
1.2 American ginseng ( <i>Panax quinquefolius</i> ) .....	1
1.3 Ginseng's active compounds: ginsenosides .....	1
1.4 <i>Pythium irregulare</i> : a pathogenic oomycete .....	4
1.5 Ginseng replant problem .....	7
1.6 Extracellular glycosidases and the metabolism of ginsenosides .....	7
1.7 Previous work .....	8
1.8 Thesis objectives .....	12
2 Materials and Methods .....	13
2.1 Cloning and expression of PiGH1-x .....	13
2.1.1 PCR .....	13
2.1.2 Cloning PiGH1-x .....	15
2.1.3 Recombinant protein expression and purification .....	15

2.1.4	Screening of colonies .....	16
2.1.5	Large-scale induction.....	17
2.1.6	Extraction under non-denaturing conditions.....	17
2.1.7	Extraction under denaturing conditions .....	17
2.1.8	6× His-tag purification.....	18
2.1.9	Buffer-exchange.....	19
2.1.10	Western-Blot analysis .....	19
2.1.11	Gel filtration chromatography.....	20
2.2	Enzyme activity assays .....	20
2.2.1	Colorimetric glycosidase assay.....	20
2.2.2	Ginsenosidase assay.....	21
3	Results and Discussion.....	23
3.1	Cloning and expression.....	23
3.1.1	PCR.....	23
3.1.2	Recombinant protein expression and purification .....	23
3.1.3	Screening of colonies .....	25
3.1.4	Large-scale protein extractions .....	26
3.1.5	His-tagged proteins .....	27
3.1.6	Non-tagged proteins.....	30
3.2	Enzymatic activity assays .....	33
3.2.1	Glycosidase activity assay .....	35

3.2.2	Ginsenosidase activity .....	36
4	General discussion .....	40
5	Conclusion and future recommendations .....	42
6	References .....	44
7	Appendices .....	50
8	Curriculum Vitae.....	62

## LIST OF TABLES

Table 1 Description and nomenclature of constructs used. ....	14
Table 2 Nomenclature and description of protein preparations for enzyme assays .....	34
Table 3 Glycosidase activity of protein preparations .....	36
Table 4 Ginsenosidase activity of protein preparations .....	38
Table 5 Summary of resulting glycosidase activity based on substrate conversion. ....	39



## LIST OF FIGURES

Figure 1 Common ginsenosides of American ginseng .....	3
Figure 2 Deglycosylation of ginsenosides mediated by extracellular ginsenosidases from <i>Pythium irregulare</i> .....	9
Figure 3 Graphic representations of the designed constructs .....	14
Figure 4 Graphic representation of the conversion of pNPG to P-nitrophenol .....	21
Figure 5 Agarose gel analysis of cloned PiGH1-x constructs. ....	24
Figure 6 Agarose gel analysis of minipreps containing PiGH1-x constructs ligated into expression vector .....	25
Figure 7 SDS PAGE analysis of small scale protein induction. Proteins expressed by <i>E. coli</i> DH5 $\alpha$ with (+) and without (-) induction with IPTG.....	26
Figure 8 SDS PAGE Analysis of large-scale expression under non-denaturing and denaturing conditions.....	27
Figure 9 SDS PAGE Analysis of Ni <sup>+2</sup> based purification of the 5'his his-tagged protein, non-denaturing conditions .....	28
Figure 10 SDS PAGE Analysis of Ni <sup>+2</sup> based purification of the 5'his and His 3 his-tagged proteins, denaturing conditions .....	29
Figure 11 Calibration UV trace using BSA and Blue Dextran. ....	31
Figure 12 UV traces and enzyme activity plots of gel filtration chromatography fractions .....	32
Figure 13 Raw EIC traces .....	37

## LIST OF ABBREVIATIONS

<b>BSA</b>	Bovine serum albumin
<b>EIC</b>	Extracted ion chromatogram
<b>G1</b>	Glycosidase 1
<b>G2</b>	Glycosidase 2
<b>G3</b>	Glycosidase 3
<b>GXVII</b>	Gypenoside XVII
<b>IMAC</b>	Immobilized metal affinity chromatography
<b>IPTG</b>	Isopropyl $\beta$ -D-1 thiogalactopyranoside
<b>LC-MS</b>	Liquid chromatography-mass spectrometry
<b>MES</b>	2-(N-morpholino)ethanesulfonic acid
<b><math>M_r</math></b>	Relative molecular weight
<b>ORF</b>	Open reading frame
<b>PBS</b>	Phosphate buffered saline
<b>PCR</b>	Polymerase chain reaction
<b>pI</b>	Isoelectric point
<b>pNP</b>	<i>p</i> -nitrophenol
<b>pNPG</b>	<i>p</i> -nitrophenyl $\beta$ -D-glucopyranoside
<b>RACE</b>	Rapid amplification of cDNA ends
<b>SDS-PAGE</b>	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
<b>TEV</b>	Tobacco etch virus

## LIST OF APPENDICES

Appendix 1 List of primers .....	50
Appendix 2 Sequence alignment of <i>full</i> .....	52
Appendix 3 Sequence alignment of 3' <i>his</i> .....	54
Appendix 4 Sequence alignment of 5' <i>his</i> .....	56
Appendix 5 Sequence alignment of <i>short</i> .....	59

# 1 INTRODUCTION

## 1.1 Ginseng overview

Ginseng is a widely studied medicinal plant that belongs to the genus *Panax* in the Araliaceae (Wen and Zimmer, 1996). It is an important part of traditional Chinese medicine, since it is considered an adaptogen, and believed to have characteristics that provide protection to the human body from physiological stress (Brekhman and Dardymov, 1969; Ang-lee et al, 2001) . Ginseng is consumed as an extracted herb (tea) or as dry ground root. There are two main types of ginseng used in traditional Chinese medicine: Asian ginseng (*Panax ginseng*) and American ginseng (*Panax quinquefolius*) (King et al., 2006). The main difference between Asian and American ginseng is in the relative concentrations of each of their active ingredients, known as ginsenosides; difference that confers them with opposite effects on the human body, while Asian ginseng is considered a stimulant, American ginseng is known for its calming effects (Schlag and McIntosh, 2006).

## 1.2 American ginseng (*Panax quinquefolius*)

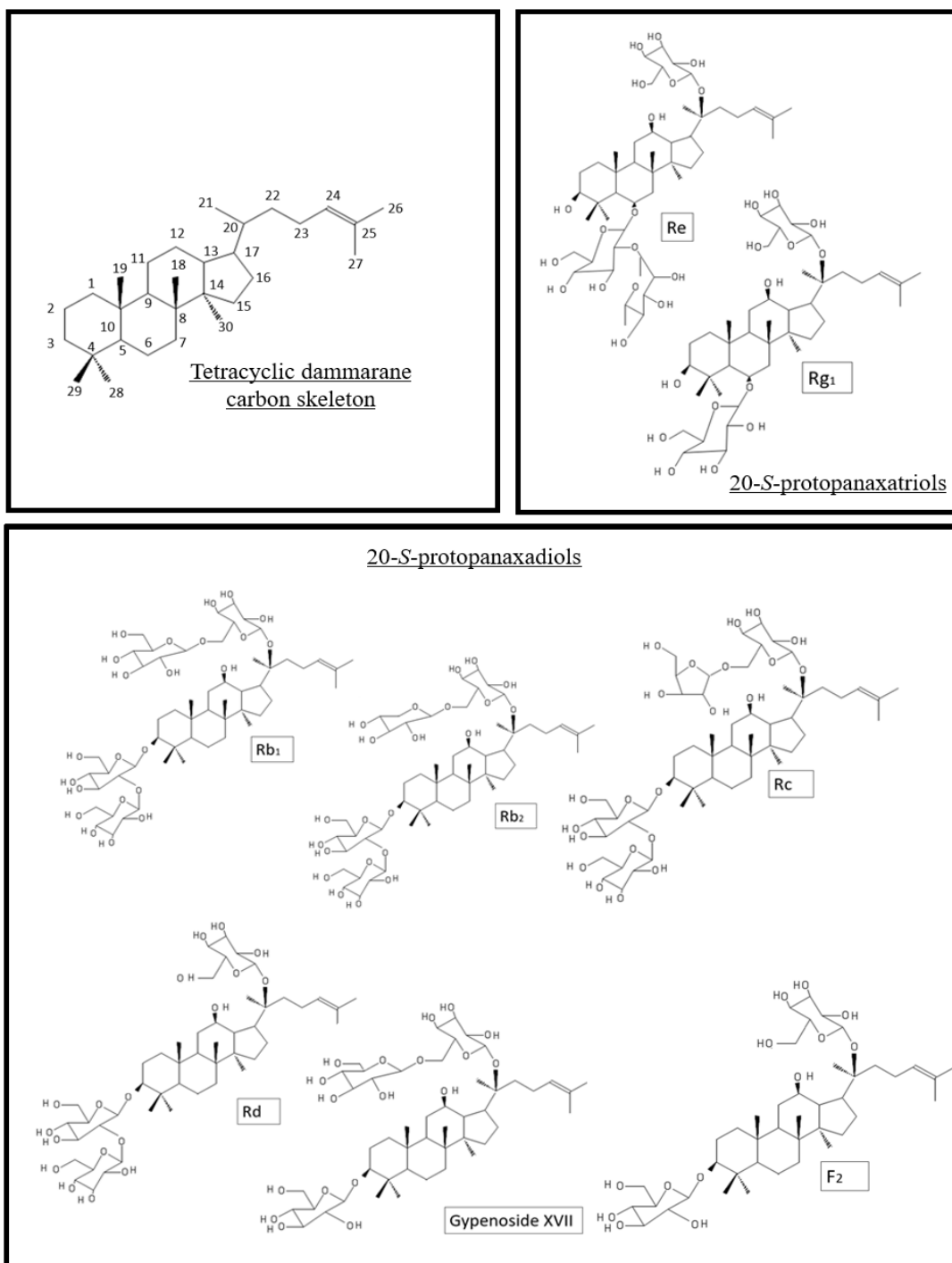
American ginseng is cultivated for commercial production mainly in Ontario, Wisconsin and British Columbia, with Ontario being the principal Canadian producer (Roy et al., 2003; Yuk et al., 2013) . Although Asian and American ginseng are the most commonly described herbal species used for their pharmacologic effects (Ang-lee et al, 2001; Qi et al., 2011; Cui et al., 2017), compared to its Asian counterpart, studies related to American ginseng constituents are limited; although there have been recent advances that have the objective of understanding the chemistry, pharmacology and function of American ginseng (Qi et al., 2011).

## 1.3 Ginseng's active compounds: ginsenosides

The pharmacological effects of ginseng have been attributed to ginsenosides, which are triterpenoid saponin glycosides (Schlag and McIntosh, 2006). Their structure is based on a tetracyclic dammarane carbon skeleton with sugars attached through hydroxyl groups at C-3 (or C-6) and C-20. Having sugars attached at both ends of the dammarane carbon

skeleton makes ginsenosides bidesmosidic, which increases their solubility, presumably affecting their in vivo function (Bernards et al., 2011). There are more than 40 known ginsenosides found in ginseng roots, the most common of which are Rb<sub>1</sub>, Rb<sub>2</sub>, Rc, and Rd (Cheng et al., 2007); these are based on a 3,12,20-trihydroxylated-20(*S*)-protopanaxadiol aglycone with sugars bound to the hydroxyl groups at positions 3 and 20. Ginsenosides Rg<sub>1</sub> and Re are also common, and are based on a 3,6,12,20-tetrahydroxylated-20(*S*)-protopanaxatriol aglycone with sugars bound to the hydroxyl groups at positions 6 and 20 (Fuzzati et al., 1999) (**Figure 1**). Ginsenosides can account for up to ~6% of the root dry weight (Schlag and McIntosh, 2006). Moreover, ginsenosides are released into the rhizosphere soil as root exudates (Yang et al., 2015).

Ginsenosides are toxic towards a few common pathogens that attack American ginseng, as well as other fungi. In vitro studies by Nicol et al. (2002) showed a clear inhibition of the growth of *Trichoderma hamatum* (a common saprotrophic fungus), and *Alternaria panax* (a foliar pathogen of ginseng). By contrast, a growth stimulatory effect was observed when some oomycete pathogens were cultured in the presence of ginsenosides. That is, ginsenosides stimulated the growth of some of the most relevant ginseng pathogens, including *Pythium irregulare*, *Phytophthora cactorum* (Nicol et al., 2003), *Ilyonectria destructans* and *Fusarium solani* (Yang et al., 2015).



**Figure 1. Common ginsenosides of American ginseng.** Ginsenoside structure is based on the tetracyclic dammarane carbon skeleton shown at the top left. Sugars, mainly glucose but also arabinose, are attached through hydroxyl groups at C-3 and C-20 of the 20-S-protopanaxadiols, and through hydroxyl groups at C-6 and C-20 of the 20-S-protopanaxatriols (structures and nomenclature based on Teng et al., 2002)

#### 1.4 *Pythium irregulare*: a pathogenic oomycete

Oomycetes (phylum Oomycota) are eukaryotic organisms, often with a filamentous growth habit. Oomycetes were initially classified as fungi because of their morphology, such as the formation of hyphae and reproduction by spores, as well as the fact that some feed on living and decaying plant material. Fungi and oomycetes comprise the majority of eukaryotic plant pathogens (Latijnhouwers et al., 2003). However, significant differences exist between fungi and oomycetes (Dean et al., 2014). Taxonomic analyses of phenotypic characteristics as well as sequence comparisons have shown that these two groups correspond to different eukaryotic branches (Latijnhouwers et al., 2003). While fungi share a common ancestor with animals, oomycetes are phylogenetically related to other protists within the Stramenopiles and their closest relatives are the heterokont golden-brown algae (Baldauf et al., 2000). Moreover, due to their biochemical differences, such as their different surface molecules that play an important role on plant recognition and defence; chemicals used to control fungi are often not effective in the inhibition of oomycetes (Hardham, 2007).

Oomycetes possess tubular hairs on the anterior flagellum of motile spores, cellulosic cell walls and are diploid during the majority of their life cycle (Hardham, 2007). Many species of the Oomycota are important plant pathogens (Dean et al., 2014). Plant pathogens can be classified as biotrophs, necrotrophs or hemibiotrophs. Biotrophs grow and reproduce in living plant tissue; obtaining nutrients by interacting with the living cells of their host, some examples are those found within the families *Albuginaceae* and *Peronosporaceae* and can cause downy mildews and white rusts. Necrotrophs kill the host tissue before colonization to feed on its dead cells; *Pythium* and *Aphanomyces* are the genera that comprise this type of oomycetes. Hemibiotrophs have intermediary pathogenic habits, establishing first a biotrophic relationship with their host, whose cells die while the infection advances; this infection pattern can be found within the genera *Phytophthora* and *Pythium* (Agrios, 2005).

The genus *Pythium* belongs to the Pythiaceae (Hendrix and Campbell, 1973), and is well known for containing soil-borne plant pathogens that can infect a large range of plants (e.g. tulip, carnation, onion, ginseng, wheat, corn, etc.) (Matsumoto et al., 2000). *Pythium* spp. primarily infect the root of the host plant, principally causing seed rot, damping off and root rot of seedlings (Martin and Loper, 1999). Mature plants can also be infected, showing root necrosis, leading to poor growth and subsequent death (Larkin et al., 1995; Martin and Loper, 1999).

The main survival structures of *Pythium* spp. are called oospores, which are thick-walled sexual spores that can resist desiccation and survive in soil for extended periods of time, without any hosts available or substrates for growth and development (Martin and Loper, 1999). These oospores show constitutive dormancy (Martin and Loper, 1999), which means that they can remain dormant presumably until they can guarantee the presence of a host (Isaac, 1992). According to Martin and Loper (1999), this type of dormancy benefits the organism by increasing the probability of survival by not allowing many propagules to germinate at the same time.

Dormant propagules can be activated by environmental changes, most likely through biochemical modifications to essential components of metabolic pathways or by changes to the membrane permeability (Isaac, 1992). Specifically, factors that can influence the transition of dormant to germinable oospores are maturity (e.g. freshly harvested vs. older oospores) (Adams, 1971), type of substrate (e.g. soil, clay, culture media, etc.) (Lumsden and Ayers, 1975; Johnson and Arroyo, 1983), CO<sub>2</sub> concentration (Johnson, 1988), humidity changes (e.g. wetting and drying) (Lifshitz and Hancock, 1984) and light conditions (Johnson, 1988).

Once activated, *Pythium* spp. produce asexual reproductive structures, named sporangia, at hyphal tips (Martin and Loper, 1999). Germination of sporangia can be induced by chemical stimulants such as amino acids, carbohydrates or other compounds exuded by the root or seeds (Lifshitz et al., 1986; Nelson, 1990; Nelson and Craft, 1989). When conditions are ideal to initiate germination of propagules, sporangia and oospores germinate by producing a germ tube. Some species also produce zoospores, which are



induced by nutrients and humidity (Martin and Loper, 1999). These zoospores do not have a cell wall; their plasma membrane acts as their outer surface (Hardam and Hyde, 1997). These structures can achieve motility due to two flagella that emerge along the ventral surface of the cell. Zoospores have specific taxic responses that guide them to specific areas of the plant (Hardham, 2007). Germinating seeds as well as roots attract them; once zoospores reach a root, they adhere, encyst and germinate to produce a penetration peg that penetrates the host cell wall in order to begin colonization by saprophytic growth of hyphae in the plant tissue (Hancock, 1977). It is hypothesized that once the pathogen has penetrated the host cell it induces changes in host cell organization and metabolism in order to obtain nutrients (Birch et al., 2006).

Specifically, *Pythium irregulare* is one of the most important ginseng pathogens (Yang et al., 2015). It is also considered one of the principal pathogenic species of its genus and has a worldwide presence (Matsumoto et al., 2000). It is known to also affect important crops such as carrot, celery, lettuce, parsnip, radish, watermelon and bean (Howard et al., 1994; Martin and Loper, 1999).

The ability of *P. irregulare* to infect ginseng is correlated with its ability to partially deglycosylate ginsenosides, presumably to generate less toxic products (Ivanov and Bernards, 2012). Indeed, the glycosidases (ginsenosidases) involved appear to be induced by the presence of ginsenosides, at least in vitro (Neculai et al., 2009). Partial deglycosylation of the 20 (S)-protopanaxadiol ginsenosides Rb1, Rb2, Rc, Rd and gypenoside XVII, via extracellular ginsenosidases, yields ginsenoside F2, which is subsequently partially assimilated by the pathogen (Neculai et al., 2009). These findings reveal that ginsenosides may act as chemoattractants or growth regulators in the pathosystem between ginseng and *P. irregulare*. Treatment of ginseng plant roots with a relatively high dose of ginsenosides resulted in a delay in infection by *P. irregulare* in pot-grown plants (Ivanov et al., 2015). Moreover, in in vitro assays, *P. irregulare* appeared to grow differently when it encountered ginsenosides, resulting in enhanced aerial mycelial growth. Therefore, it has been hypothesized that ginsenosides may cause a modification in *P. irregulare*'s growth habit, which may ultimately lead to the build-up of inoculum, and in this way facilitate infection (Ivanov et al., 2015).

## 1.5 Ginseng replant problem

The elucidation of the role of ginsenosides in the infection mechanism by *P. irregulare* is of importance, since this interaction may be related to a major issue in commercial ginseng production, known as the “replant problem”. That is, new ginseng crops planted in previous ginseng gardens, suffer a yield reduction because of a lack of good seedling growth and germination, as well as the appearance of severe diseases (Yang et al., 2015). This is presumed to be caused by the accumulation of pathogens in previously harvested soil (Kernaghan et al., 2007).

In studies performed in a different species of ginseng (*Panax notoginseng*, also known as Sanqi ginseng), it was hypothesized that the secondary metabolites released to the soil by the plant could accumulate and become toxic to new ginseng seedlings replanted in the same soil. Since Sanqi ginseng contains at least 20 different ginsenosides, which contribute to around 6% of its dry mass, they could be a factor in ginseng replant failure (Yang et al., 2015). Seedling emergence as well as its survival rate decreased significantly with continuous planting over 3 years. Moreover, ginsenosides were identified in the roots and soil, and specifically, ginsenosides R<sub>1</sub>, Rg<sub>1</sub>, Re, Rg<sub>2</sub> and Rd were demonstrated to be toxic impeding seedling emergence, growth and diminishing root cell vigor (Yang et al., 2015). Replant disease is complex, and likely involves many factors. Nevertheless, any factor than can contribute to the build-up of inoculum in the soil can also potentially contribute to the replant problem. The influence on pathogen growth of ginsenosides present in ginseng garden soil cannot be ruled out as a contributing factor.

## 1.6 Extracellular glycosidases and the metabolism of ginsenosides

Many plant species produce antifungal saponins, which are used as a barrier to prevent infection by pathogenic fungi (Bowyer et al., 1995). However, some fungal pathogens can degrade/metabolise/detoxify such host plant saponins via extracellular enzymes. As an

example, *Gaeumannomyces graminis*, which is a fungus that commonly infects cereals, can successfully infect a host containing saponins only if it can produce an extracellular saponinase enzyme, denominated avenacinase. Therefore it is suggested that the “saponin-saponinase” interactions are important to the infection mechanism of some pathogens (Bowyer et al., 1995).

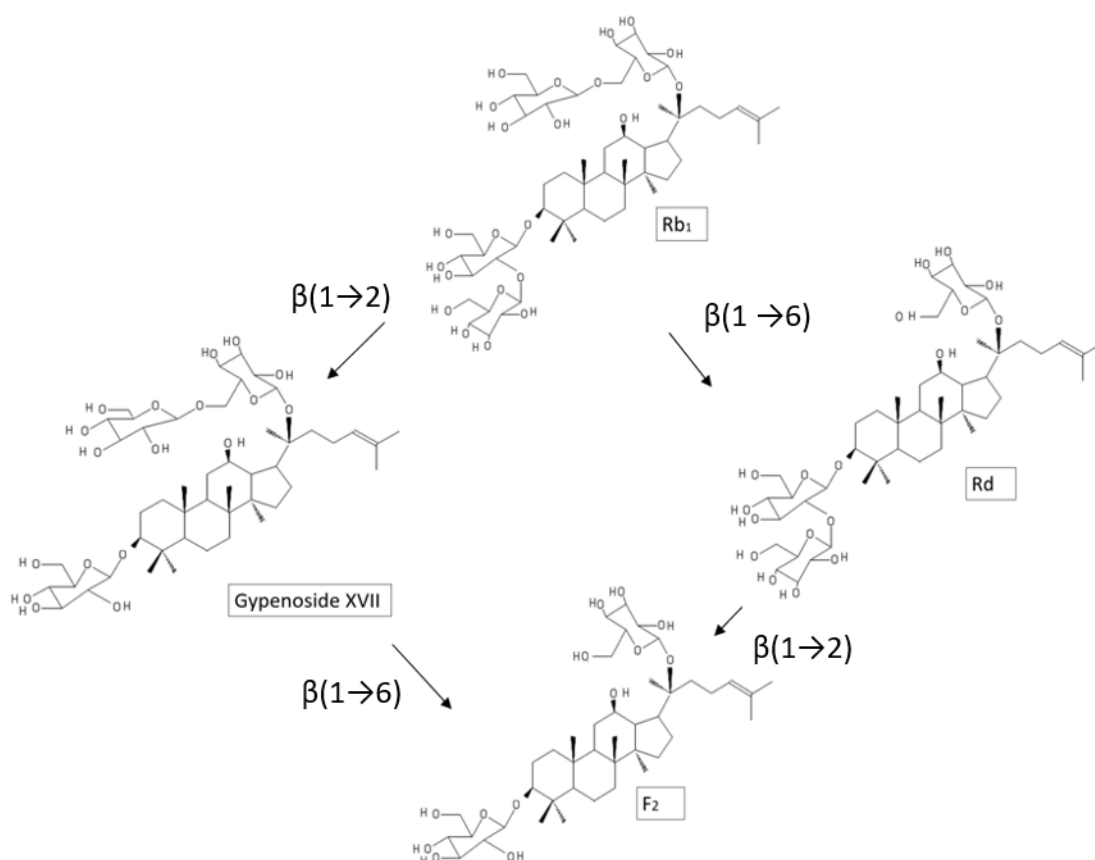
In the case of *P. irregulare*, it has been hypothesized that its pathogenicity varies significantly between different hosts (Harvey et al., 2001). This suggests that different plants secrete different secondary metabolites to modify the interaction between the pathogen and the host. Nonetheless, the role of extracellular enzymes in oomycete infection has not been fully established yet (Hardham, 2007).

Furthermore, ginsenosides are known to be mildly fungitoxic but seem to promote the growth of root pathogens (g.e. *P. irregulare* and *Ilyonectria destructans*), this happens via partial metabolism of ginsenosides using extracellular saponinases (g.e. ginsenosidases). Specifically, *P. irregulare* seems to prefer 20(*S*)-protopanaxadiol ginsenosides, partially metabolizing them and yielding ginsenoside F2, while it does not have the ability to metabolize any of the 20(*S*)-protopanaxatriol ginsenosides (Yousef and Bernards, 2006).

## 1.7 Previous work

In previous studies in the Bernards Lab (University of Western Ontario, London, ON, Canada), the metabolism of ginsenosides by *P. irregulare* during growth, in Czapek-Dox minimal medium, was monitored in a time-course manner. Ginsenosides were collected over the course of 9 days, and analyzed by liquid chromatography. As time progressed, all major protopanaxadiol ginsenosides (i.e., Rb1, Rd and GXVII) decreased, while a new compound, later identified as ginsenoside F2, accumulated (Yousef and Bernards, 2006). At the same time, it was observed that low concentration of protein was secreted into the culture medium by *P. irregulare*. The secreted proteins were fractionated using gel filtration and tested for glycosidase activity. Increasing glycosidase activity was measured in two distinct fractions. In contrast, *P. irregulare* grown in the absence of ginsenosides displayed little extracellular glycosidase activity. Subsequent purification of the secreted glycosidases yielded two distinct activity peaks. The fractions showing

glycosidase activity were then tested for ginsenosidase activity by incubation with a crude mixture of ginsenosides and were shown to deglycosylate ginsenosides. One purified fraction was enriched in  $\beta(1\rightarrow2)$  glycosidase activity, while the another was enriched in  $\beta(1\rightarrow6)$  activity. In other words, the fractions that corresponded to the first activity peak had the ability to deglycosylate ginsenoside Rb1, converting it into gypenoside XVII, as well as converting Rd into F2. The second activity peak, on the other hand, metabolized Rb1 into Rd, and gypenoside XVII into F2 (**Figure 2**) (Neculai et al., 2009).



**Figure 2** Deglycosylation of ginsenosides mediated by extracellular ginsenosidases from *Pythium irregulare*. This schematic representation shows the conversion between ginsenosides caused by extracellular ginsenosidases (structures and nomenclatures based on Yousef and Bernards, 2006).

Altogether, three glycosyl hydrolases with ginsenosidase activity were purified from *P. irregulare* culture filtrates. The two distinct glycosidase activities observed were categorized as glycosidase 1 (G1) and glycosidase 2 (G2). Electrophoretic analysis showed a clear protein band for G1 and two very close bands for G2, suggesting that this fraction contains a mixture of two enzymes, denoted as G2/G3 (Neculai et al., 2009).

SDS-PAGE and gel filtration were used to estimate the molecular weight of the isolated enzymes. G1 is a homodimer with two 78 kDa subunits, while G2 and G3 were predicted to be 61 and 57 kDa, respectively. Molecular weight ( $M_r$ ) and Isoelectric Point (pI) values of the purified enzymes were found to be on the same range as fungal saponin-detoxifying glycosidases such as tomatinase (Quidde et al., 1998) and avenacinase (Quidde et al., 1999) of *Botrytis cinerea*. However, they were found to be closer in size to ginsenoside-metabolizing enzymes from bacteria and plants (i.e. *Panax ginseng*, *Bifidobacterium breve* and *Fusobacterium*) (Neculai et al., 2009).

Purified proteins were excised from gels, digested with trypsin and subjected to LC-MS/MS-based sequencing. *De-novo* amino-acid sequence data obtained from ginsenosidases G2 and G3 yielded peptides with conserved amino acid regions with 100% confidence scores for six proteins annotated as  $\beta$ -glycosidases from the *P. ultimum* Genome Database. These sequences were then used to design degenerate primers with the objective of amplifying the genes that encode these enzymes, using total cell DNA as well as RNA as templates (Ivanov, 2015). This information facilitated the prediction of the 780 bp size amplicon obtained by PCR using *P. irregulare* genomic DNA as template and the degenerate primers noted above (Ivanov, 2015).

To determine the full coding cDNA sequences for the potential glycosidases, RACE (Rapid Amplification for cDNA Ends) experiments were performed (Ivanov, 2015). Both 5'RACE and 3'RACE resulted in amplicons of approximate sizes of 600 bp and 1100 bp, respectively; when aligned, the full coding sequence of one putative  $\beta$ -glycosidase (1767 bp without introns) was revealed. The predicted protein encoded by the obtained sequence, is expected to contain 589 amino acids and has a predicted molecular weight of 66.14 kDa and a pI of 6.25 (Ivanov, 2015).

Homology modeling revealed that the *P. irregularis* sequence most closely resembled a glycosyl hydrolase family 1 type  $\beta$ -glycosidase from wheat (Ivanov, 2015). The 3D model, generated using the Phyre2.0 website and rendered using the Chimera 1.10 software demonstrated that the predicted protein structure contains glutamine residues (Glu 198 and Glu 416) that act as a potential active site where glucose residues may bind. The putative  $\beta$ -glycosidase from *P. irregularis* was denoted PiGH1-x, where the “x” is a place-holder pending confirmation of catalytic activity.

## 1.8 Thesis objectives

Extracellular glycosidases produced by *P. irregulare* are clearly associated with the ability of this organism to infect ginseng. Nevertheless, the specific mechanism remains unclear, which is why it is important to further analyse the function of these enzymes in the infection mechanism. Therefore, the objectives of the present thesis are:

1. Cloning and expression of PiGH1-x produced by *P. irregulare*.
2. Characterisation of the recombinant PiGH1-x protein with respect to its ability to deglycosylate ginsenosides.

## 2 MATERIALS AND METHODS

All the reagents used in this project were obtained from Sigma-Aldrich, except when noted otherwise.

### 2.1 Cloning and expression of PiGH1-x

#### 2.1.1 PCR

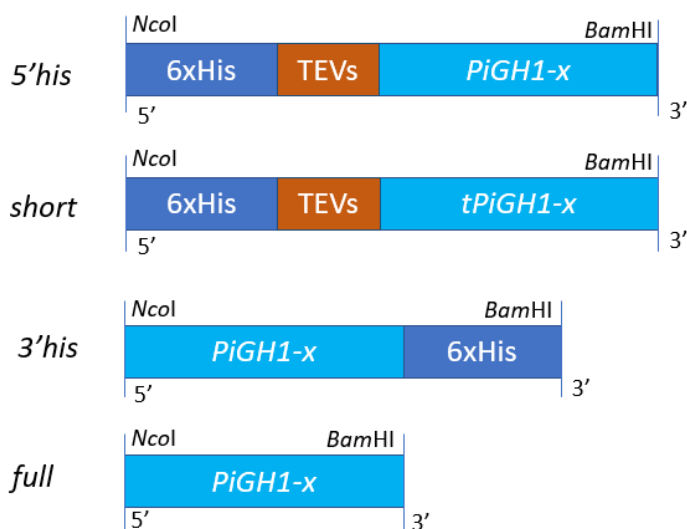
The Polymerase Chain Reaction (PCR) was used in order to amplify the full Open Reading Frame (ORF) of *PiGH1-x* from the stock of a clone generated by Dimitre Ivanov (Ivanov, 2015); which was used as a template. New primers were designed to generate 4 different constructs for protein expression (**Table 1, Figure 3**), based on the protocol by Rahman et al. (2011). A detailed list of primers can be found in **Appendix 1**.

The *PiGH1-x* constructs were amplified using Platinum Taq DNA Polymerase High Fidelity (ThermoFisher Scientific) to ensure high specificity. The master mix used was composed of 10X High Fidelity PCR buffer, 50 mM MgSO<sub>4</sub>, 10 mM dNTP mix, 10 µM forward primer and 10 µM reverse primer. The PCR program was as follows. The pre-denaturation stage was first run at 94°C for 3 min. The next cycle was repeated during 35 times and consisted of 3 steps: the denaturing step at 94°C for 45s, the annealing step at 54°C for 45s and the extension at 72°C for 2:30 min. Final elongation was run at 72°C for 10 min, and the final hold was set at 4°C.



**Table 1 Description and nomenclature of constructs used.**

CONSTRUCT NAME	EXPECTED AMPLICON
<i>5'his</i> Size: 1839 bp	Full <i>PiGH1-x</i> sequence, with a TEV recognition site and a 6× His tag attached to the 5' end.
<i>short</i> Size: 1818 bp	Truncated <i>PiGH1-x</i> (t <i>PiGH1-x</i> ) sequence (signal peptide sequence removed) with a TEV recognition site followed by a 6× His tag attached to the 5' end.
<i>3'his</i> Size: 1819 bp	Full <i>PiGH1-x</i> sequence, with a 6× His tag attached to the 3' end.
<i>full</i> Size 1785 bp	Full <i>PiGH1-x</i> sequence, without any purification tags.

**Figure 3 Graphic representations of the designed constructs.** ORF of the protein is shown in light blue, 6×His tag sequence in dark blue and the TEV recognition site in orange. All sequences are flanked by *NcoI* and *BamHI* restriction sites.

### 2.1.2 Cloning *PiGH1-x*

PCR amplification described in section 2.1.1 resulted in amplicons of roughly 1.8 kb (corresponding to the full length clone). Bands of the expected size were extracted from 1% agarose gels using the QIAquick Gel Extraction Kit (Qiagen N.V.) followed by cleanup using the QIAquick PCR Purification Kit (Qiagen N.V.), in order to completely remove any residual primers. Each of the resulting fragments were ligated into a cloning vector pGEM-T Easy, according to the instructions of the manufacturer (Promega Corp. Fitchburg, WI, USA). Competent *E. coli* DH5 $\alpha$  cells were then transformed by heat-shock, per the manufacturer's instructions (Invitrogen), to incorporate the plasmids containing the desired constructs.

After ampicillin-based selection of transformed *E. coli*, DNA was extracted using the Qiaprep Spin Mini-Prep Kit (Qiagen N.V.). To verify that the correct constructs were inserted into the vector, resulting DNA was analyzed using T7 and Sp6 primers in an Applied Biosystems 3730 Analyzer at the DNA Sequencing Facility at the London Regional Genomics Centre (Robarts Research Institute, London, ON, Canada). Sequence verification was performed by aligning the resulting sequences with the expected ORF, and sequences for restriction sites and purification tags where applicable, using the software DNAMAN Version 6 (Lynnon Corporation).

### 2.1.3 Recombinant protein expression and purification

Molecular cloning, expression and purification of recombinant PiGH1-x was initially performed in collaboration with Dr. John McCormick (Microbiology and Immunology, UWO) using his standard protocol (McCormick and Schlievert, 2003). Briefly, the protein genes were cloned into the pET28a plasmid, using PCR to incorporate *Nco*I and *Bam*HI restriction sites through the forward and reverse primers, respectively. The *Nco*I was engineered in a way that allowed an in frame translational fusion with the ATG start codon of the target construct (McCormick and Schlievert, 2003). Competent *E. coli* BL21(DE3) cells were transformed by heat shock, according to manufacturer's

instructions (Novagen), and 4 clones were generated, each of them containing one of the desired expression cassettes. This *E.coli* strain carries the DE3 lysogen that encodes the T7 RNA polymerase under inducible control by  $\beta$ -D-galactopyranoside (IPTG) (McCormick and Schlievert, 2003).

#### 2.1.4 Screening of colonies

Colonies containing the plasmids of interest were screened by a small-scale induction to verify their capability of expressing the desired recombinant protein. For this purpose, cultures were grown overnight in 4 mL of LB medium supplemented with 50  $\mu$ g/mL of Kanamycin. Next, 5  $\mu$ L of the resulting overnight culture were then subcultured into 5 mL of LB medium with 50  $\mu$ g/mL of Kanamycin.

Diluted cultures were grown at 37°C on a shaker at 300 rpm for 2.5 h prior to induction using 200  $\mu$ M IPTG and kept in the incubator overnight. For each construct tested, both an induced and an uninduced culture were grown.

The overnight screening cultures were pelleted at 6000  $\times g$  for 3 minutes using a tabletop centrifuge, and the supernatant was discarded. Next, cells were resuspended in 1 mL Hepes buffer (10 mM Hepes, 150 mM NaCl, pH 7.4) and sonicated using a Branson Sonifier 450 (30 pulses at output 4 and duty 40%). Samples were then centrifuged in a benchtop microcentrifuge at 13000  $\times g$  for 2 min (McCormick and Schlievert, 2003).

The supernatants, containing crude protein extracts, were analyzed on 12% SDS-PAGE gels, which were run for 30 min at 40 V followed by 1 h at 100 V in a Bio-Rad electrophoresis chamber, according to manufacturer's instructions. Gels were then stained in a staining solution containing 1 g of Coomassie Brilliant Blue dissolved in 1 L of destain solution (20% methanol, 10% glacial acetic acid, 70% ddH<sub>2</sub>O), with light shaking for 1 h. Afterwards, the staining solution was removed and gels were soaked in destain solution for 2 h to allow the existing protein bands to be visualized.

### 2.1.5 Large-scale induction

Successfully screened clones were subcultured at a 1:1000 dilution into 300 mL of LB medium with 50  $\mu\text{g/mL}$  of Kanamycin and grown at 37°C with shaking at 300 rpm for 2.5 h. Protein expression was then induced by the addition of 200  $\mu\text{M}$  IPTG. Cultures were grown overnight prior to protein isolation (McCormick and Schlievert, 2003).

### 2.1.6 Extraction under non-denaturing conditions

Large overnight induced cultures (resulting from *full*, *5'his*, and *3'his*) were centrifuged in a large centrifuge bottle, using a JLA 16.250 rotor at  $8000 \times g$ , and the pellet was resuspended in 10 mL of Hepes buffer with 10  $\mu\text{L}$  of DNase I (Thermo Scientific, 10 mg/mL) and 50  $\mu\text{L}$  of chicken egg lysozyme (Sigma, 10 mg/mL). Samples were incubated for 1 h at 4°C. An uninduced control was also prepared for each sample, following the same method but without addition of IPTG.

Next, the samples were sonicated for 3 sets using a Branson Sonifier 450 (each set consisting of 50 pulses at output 4 and duty 40%). Cellular debris was removed by centrifugation for 10 minutes at  $12000 \times g$  (using a JA-25.50 rotor) and the supernatant was collected as it corresponded to the crude protein extract. Extracts were analyzed by SDS-PAGE using 12% gels.

### 2.1.7 Extraction under denaturing conditions

In order to avoid protein aggregation, recombinant versions of PiGH1-x (expressed from constructs *full*, *5'his* and *3'his* using the method described in section 2.1.1) were extracted under denaturing conditions, following the “Urea protocol” provided by the Instruction Manual of the Profinity™ IMAC Resin (Bio-Rad). An uninduced control was also prepared for each sample, following the same protocol, but without addition of IPTG.

Briefly, large overnight induced cultures were centrifuged in a large centrifuge bottle, using a JLA 16.250 rotor at  $8000 \times g$ , and the pellet was resuspended in 10 mL of Lysis Buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, Urea 6 M, and 1 mM PMSF). After 1 hour of incubation at 4°C the cells were sonicated for 4 sets (each set

consisting of 30 pulses at output 4 and duty 40%). The cell suspension was centrifuged at  $12000 \times g$  for 10 minutes using a JA-25.50 rotor, and the pellet was resuspended in 10 mL of PBS buffer (8M urea, pH 7.5). Protein presence was verified through electrophoresis using 12% SDS-PAGE gel.

## 2.1.8 6× His-tag Purification

### 2.1.8.1 Non-denatured protein extracts

The recombinant tagged proteins expressed in *E. coli* BL21 (resulting from the constructs denominated: 5'his, short, and 3'his), were passed through a column with Profinity<sup>TM</sup> IMAC Resin (Bio-Rad) charged with Ni<sup>2+</sup> to bind the 6× His-tag. Elution was performed using an imidazole gradient from 0-200 mM (Günther et al., 2007).

The fraction with the highest concentration of imidazole was treated with a recombinant auto-inactivation-resistant His<sub>7</sub>::TEV protease to cleave the His<sub>6</sub>::rPiGH1-x and His<sub>6</sub>::rtPiGH1-x fusion proteins (Rahman et al., 2011). The aim of this method was to separate the purification tag as well as the His<sub>7</sub>::TEV protease from rPiGH1-x and rtPiGH1-x with the use of a second nickel affinity column (McCormick and Schlievert, 2003). Recombinant protein purity was assessed using a 12% SDS-PAGE gel and protein yield quantified using the Quick Start<sup>TM</sup> Bradford Protein Assay kit.

### 2.1.8.2 Denatured protein extracts

The recombinant tagged proteins expressed in *E. coli* BL21, resulting from the constructs denominated 5'his, and 3'his (as there was not a difference in protein yields between constructs, the clone containing the construct short was not used in subsequent analyses to avoid activity issues), and extracted using denaturing conditions were passed through a column with Profinity<sup>TM</sup> IMAC Resin (Bio-Rad) charged with Ni<sup>2+</sup>, using denaturing conditions and following a protocol for on-column refolding as per manual instructions. This protocol includes a re-folding step, and is expected to yield a re-natured protein upon final elution.

### 2.1.9 Buffer-exchange

The protein extracts obtained from the extraction described in 2.1.7, as well as the non-refolded protein from the flow-through samples obtained from method 2.1.8.2, were buffer exchanged overnight (i.e. 16 hours), against three changes of 1 L of 10 mM MES buffer (pH 6.5), through dialysis using a molecular porous membrane tubing (Spectra/Por, 12-14 kDa molecular weight cut-off).

### 2.1.10 Western-Blot Analysis

Western-Blot Analysis was performed in order to detect the affinity tag expressed in the protein preparations 3'his and 5'his. This was performed according to the “General Protocol for Western Blotting” provided by Bio-Rad, using a conjugated Anti-his (C-Term)-HRP Antibody (Novex).

First, a 12% SDS gel was run in order to separate the protein fractions of the samples (running parameters as described in previous sections, without the staining process). Next, the protein gel was transferred to a nitrocellulose membrane with a pore size of 0.45  $\mu$ m (Bio-Rad), which was placed on the running cassette and into the transfer tank. Transfer was performed in the cold room overnight at a constant current of 10 mA, according to the manufacturer's instructions.

Afterwards, the blot was rinsed and stained with Ponceau S staining buffer (0.2% (w/v) Ponceau S, 5% glacial acetic acid), rinsed again and followed by incubation with the conjugated Anti-his (C-Term)-HRP Antibody in 5% non-fat dried milk, for 3 h in the cold room.

Blots were visualized using the Amersham ECL Western Blotting Detection Reagent (GE Healthcare). This was achieved following the manufacturer's instructions, mixing an equal volume of detection solution 1 and detection solution 2, making sure that the total volume covered the membrane. Next, the excess buffer was drained and the blot was wrapped using plastic wrap. Then, the wrapped blot was placed in an X-ray film cassette and a sheet of autoradiography film (Fujifilm, Fuji Medical X-Ray Film, 13×18 cm) was placed on top of the membrane. The cassette was exposed for 5 min.

The film was developed in the dark room by soaking the film in developing solution (Kodak Carestream GBX) for 15 min, followed by submersion in fixer solution for 5 min (Kodak Carestream GBX). The film was then rinsed with distilled water and allowed to air-dry.

### 2.1.11 Gel filtration chromatography

After extraction following the protocol mentioned in sections 2.1.6 and 2.1.7, the recombinant protein extracts (non-denatured and renatured) resulting from the expression cassette *full*, were partially purified using gel filtration chromatography on a Sephacryl S-200 HR column (GE Healthcare), using 10 mM MES buffer (pH 6.5, containing 100 mM KCl) as the mobile phase at a flow rate of 0.5 ml/min. Collection of 1 mL fractions began after one hour of elution.

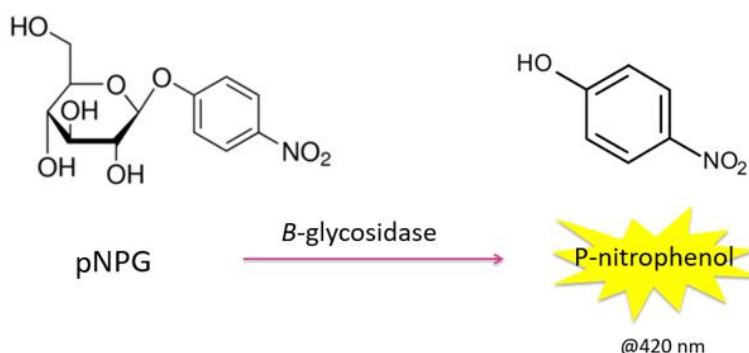
A preparation containing 1 mg/mL each of Blue Dextran and BSA (bovine serum albumin) was also run under the same conditions and used as molecular weight markers. Blue Dextran has a high mass ( $\sim 2 \times 10^6$  kDa) and is used to track the void volume of the column, while BSA has a molecular mass is of 66.4 kDa (Sigma Aldrich.), and was used to determine the expected elution volume for recombinant PiGH1-x.

## 2.2 Enzyme activity assays

### 2.2.1 Colorimetric Glycosidase Assay

Protein extracts and fractions obtained from the gel filtration column, were tested for glycosidase activity using a colorimetric method based on the use of *p*-nitrophenyl  $\beta$ -D-glucopyranoside (pNPG) as substrate. The enzyme preparations were incubated with pNPG at 30°C overnight (reaction mixture contained 90  $\mu$ L of enzyme preparation in 10 mM MES buffer pH 6.5 and 10  $\mu$ L of 50 mM pNPG). The reaction was stopped with the addition of 100  $\mu$ L of 1 M Na<sub>2</sub>CO<sub>3</sub> to enhance color formation. The reaction product, *p*-nitrophenol (pNP), which yielded a yellow solution, was measured at 420 nm **Figure 4** (Neculai et al., 2009). Successive fractions indicating an activity peak were pooled together. Negative controls (90  $\mu$ L of 10 mM MES buffer pH 6.5 and 10  $\mu$ L of 50 mM pNPG) were used to determine background absorbance at 420 nm.

For each gel filtration column run, the pooled protein contained in the 10 fractions defining the void volume, as well as the 10 fractions representing the ~ 66 kDa protein elution volume were concentrated using an Amicon Ultra-15 Centrifugal Filter Unit (Millipore). The concentrated samples were subsequently used for enzyme assay using pNPG and ginsenosides. Similarly, the two His-labeled proteins (3'his and 5'his), were assayed before and after application to a Ni-column. Enzyme specific activity was measured in nanomoles of PNP/h/μg of protein preparation.



**Figure 4 Graphic representation of the conversion of pNPG to P-nitrophenol.**

Reaction used to assay for glycosidase activity, when positive the reaction product emits a yellow colour which can be measured using spectrophotometry at 420 nm (reaction based on description by Neculai (2009)).

### 2.2.2 Ginsenosidase Assay

Ginsenosidase activity of induced and uninduced protein preparations, was tested using the protopanaxadiol ginsenoside Rb1 as substrate. Enzymatic reaction mixtures containing 10 μL of the substrate (2 mg/mL) and 90 μL of the protein solution in 10 mM MES buffer (pH 6.5), were incubated at 30°C overnight. The reaction was stopped with the addition of 100 μL of methanol, and the samples centrifuged at top speed in a microcentrifuge for 5 minutes. The reaction products were transferred to LC vials and analyzed using High Performance Liquid Chromatography coupled with Mass Spectrometry (LC-MS) (Ivanov et al., 2015). For data analysis, extracted ion chromatograms were generated using the expected exact mass of Rb1 ( $[M+Na^+]^+$  at



1131.5922 m/z, retention time: 7.7 min), Rd ( $[M+Na]^+$  at 969.5393 m/z, retention time: 8.4 min) and Gypenoside XVII ( $[M+Na]^+$  at 969.5393 m/z, retention time: 8.6 min). The  $Na^+$  adducts were generated by adding Na-Acetate (1 mg/L) to the mobile phase. Ginsenosidase specific activity was measured in nmol of product/h/mg of protein preparation.

## 3 RESULTS AND DISCUSSION

### 3.1 Cloning and expression

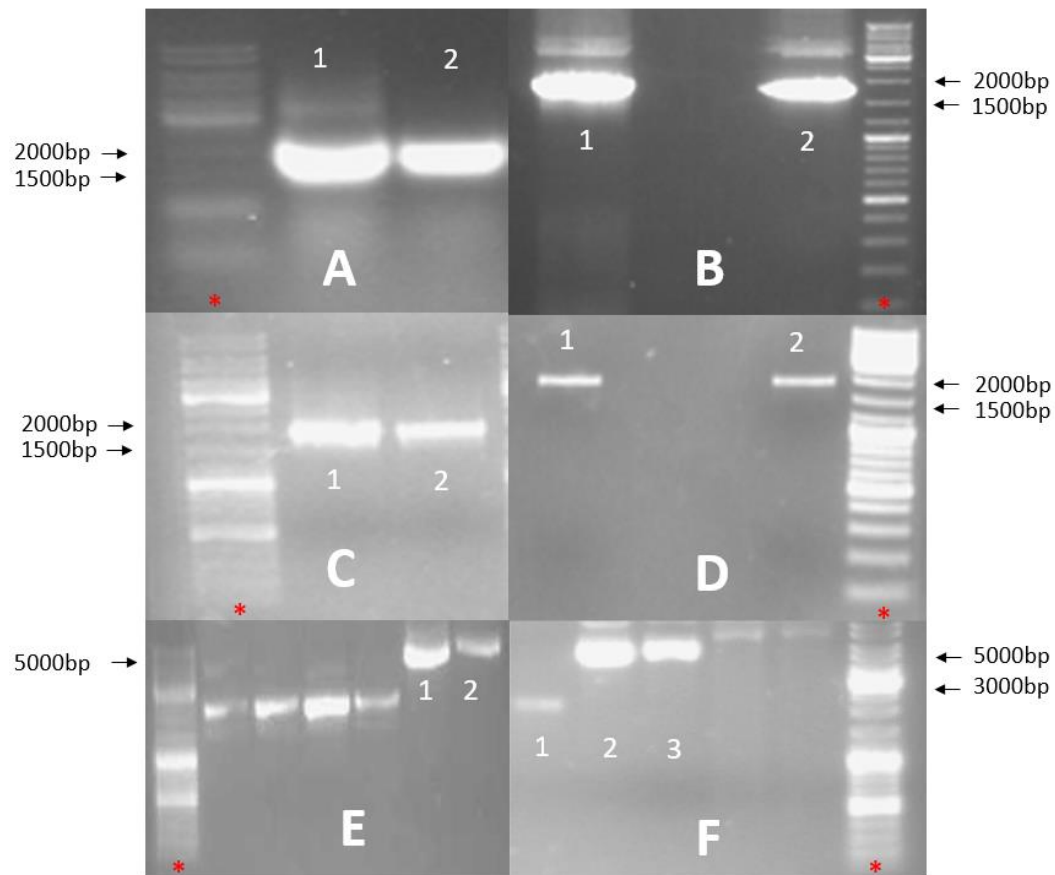
#### 3.1.1 PCR

The desired *PiGHI-x* constructs (i.e. *full*, *3'his*, *5'his* and *short*), were successfully amplified by PCR, cloned into the pGEM-T easy vector, and transferred into *E. coli* DH5 $\alpha$  (**Figure 5**). DNA minipreps from *E.coli* DH5 $\alpha$  clones were prepared, and the clones sequenced for verification (**Appendices 2-5**). Alignment of the resulting sequences with the original *PiGHI-x* sequence, confirmed that the constructs contained the desired sequences with 100% identity to the original ORF of the protein of interest, as well as the desired restriction sites and tags added for each construct.

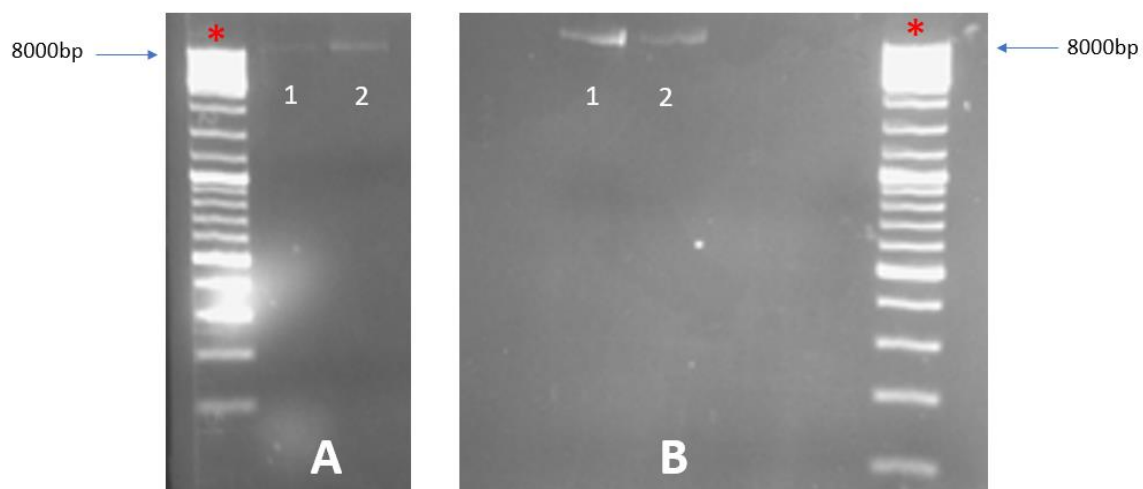
#### 3.1.2 Recombinant protein expression and purification

##### 3.1.2.1 Molecular cloning into expression vector

Fragments were extracted from verified clones using the *Nco*I and *Bam*HI restriction sites, ligated into the expression vector pET28a, and transferred into *E.coli* BL21 (DE3) cells. DNA minipreps prepared from each of the designed constructs yielded PCR product bands of the correct size (**Figure 6**).



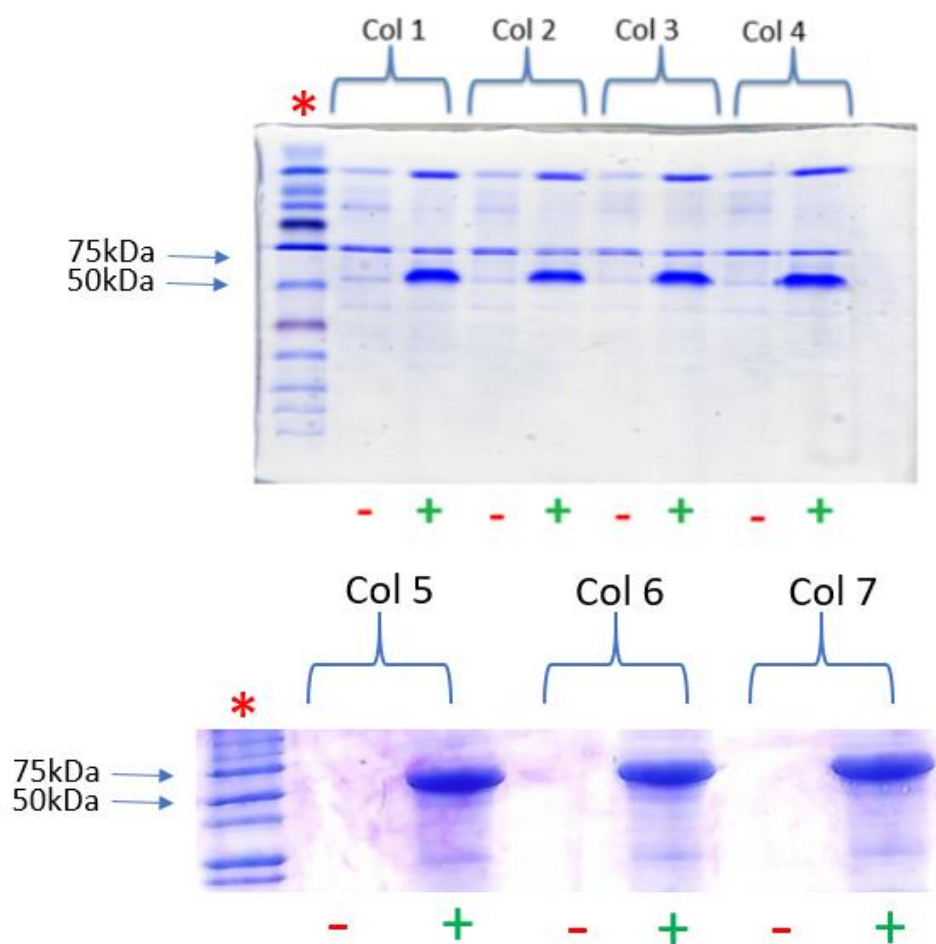
**Figure 5. Agarose gel analysis of cloned *PiGHI-x* constructs.** A) PCR product of the *full* (1) and *3'his* (2) clones. B) PCR product of the *5'his* (1) and *short* (2) clones. C) Gel extracted fragments of *full* (1) and *3'his* (2) constructs. D) Gel extracted fragments of *5'his* (1) and *short* (2) constructs. E) Inserts *full* (1) and *3'his* (2) ligated into the MCS of pGEM-T easy. F) Empty vector pGEM-T easy (1), inserts *5'his* (2) and *short* (3) ligated into the MCS of pGEM-T easy. Ladders are marked with an asterisk.



**Figure 6 Agarose Gel Analysis of minipreps containing *PiGHI-x* constructs ligated into expression vector.** A) Constructs *full* (1) and *3'his* (2) ligated into the MCS of the pET28a expression vector. B) Constructs *5'his* (1) and *short* (2) ligated into the MCS of the pET28a expression vector. Ladders are marked with an asterisk.

### 3.1.3 Screening of colonies

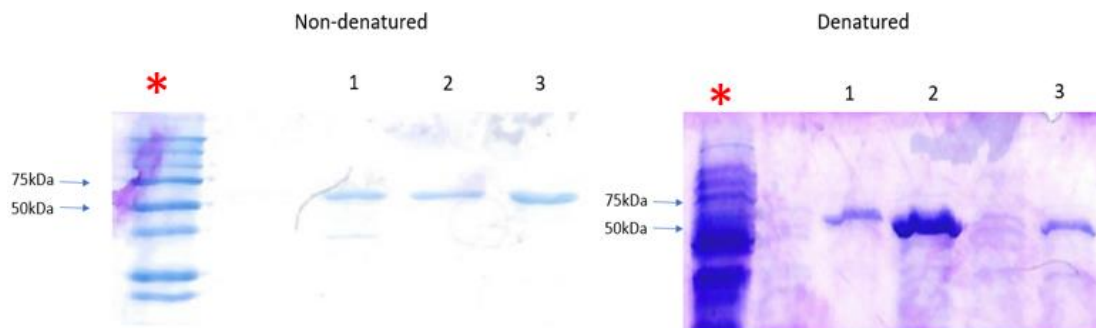
Screening of colonies induced with IPTG for target protein expression demonstrated that only colonies harbouring *PiGHI-x* constructs expressed a protein of the expected size of around 66.1 kDa (**Figure 7**). There was no apparent difference in the expression levels among the constructs (i.e., with and without the signal peptide removed), so, in order to minimize the number of proteins to process and avoid potential activity issues associated with a truncated protein, the *short* clone was not used for subsequent experiments.



**Figure 7 SDS PAGE Analysis of small scale protein induction.** Proteins expressed by *E. coli* DH5 $\alpha$  with (+) and without (-) induction with IPTG. Colonies 1-2, protein expressed from the *5'his* construct. Colonies 3-4, protein expressed from the *short* construct. Colony 5, protein expressed from the *5'his* construct. Colony 6, protein expressed from the *3'his* construct. Colony 7, protein expressed from the *full* construct.

### 3.1.4 Large-scale protein extractions

Large scale extractions were performed using proteins expressed from *5'his*, *3'his* and *full* constructs. Both IPTG-induced and non-induced protein extracts (used as controls) were prepared. Protein extraction yielded clear protein bands of the expected size (**Figure 8**).



**Figure 8 SDS PAGE Analysis of large-scale expression under non-denaturing and denaturing conditions.** 1) 3'his 2) Full 3) 5'his. Ladders are marked with an asterisk.

### 3.1.5 His-tagged proteins

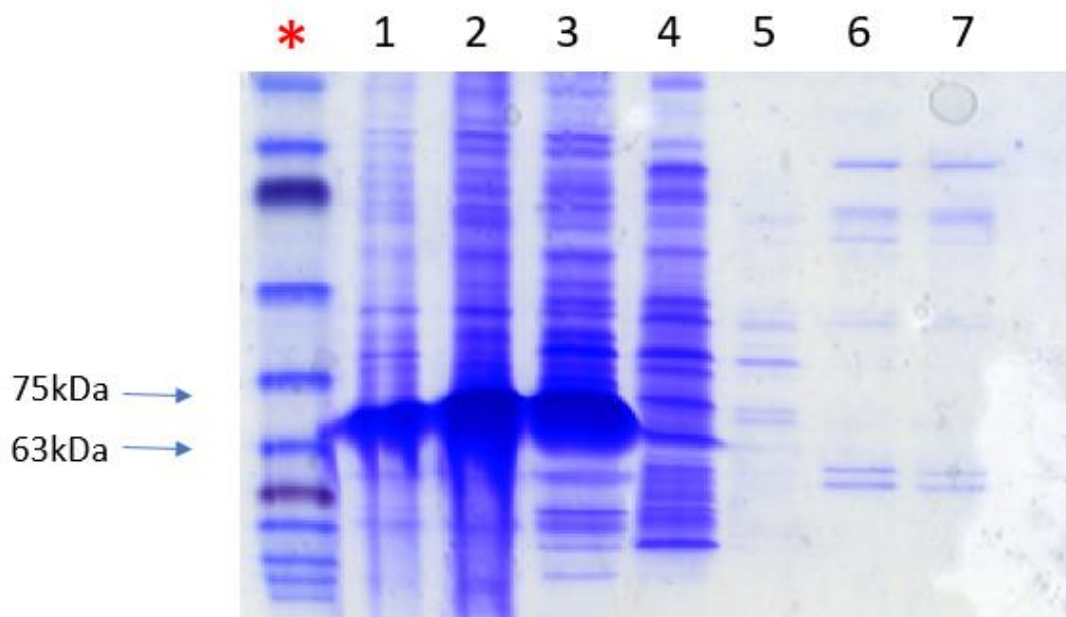
#### 3.1.5.1 Western-Blot analysis

Western Blot analysis was performed on protein extracts expressed from constructs 3'his and 5'his (as those are the proteins that contained a 6× His affinity tag) and little or no signal was detected using the anti-His antibody (data not shown). A positive control (StFAωH1-6×His) yielded a strong band, confirming the function of the antibody (data not shown). The lack of signal from recombinant PiGH1-x was unexpected, but may have resulted from the tag being inaccessible to the antibody in the folded protein. It has been reported that even under denaturing conditions, the folding of a protein can make a portion of the tag inaccessible (Debeljak et al., 2006).

#### 3.1.5.2 Nickel column purification

Despite the lack of signal from His-tagged recombinant PiGH1-x in Western Blots, extracts were passed through a Ni-affinity column. However, when non-denatured protein extracts (i.e. 5'his and 3'his) were passed through a nickel affinity column for purification, a large proportion of the recombinant protein was recovered in the flow-through and wash fractions (**Figure 9**). Presumably, this was due to the 6× His-tag not being exposed in a way that it could interact with the nickel in the column. Bornhorst and Falke (2000) suggested that moving the affinity tag to the opposite end of the protein could improve the ability of the protein to bind to the column. Nonetheless, although both

5'his (**Figure 9**) and 3'his (data not shown) were his-tagged each on an opposite end, column binding was unsuccessful.



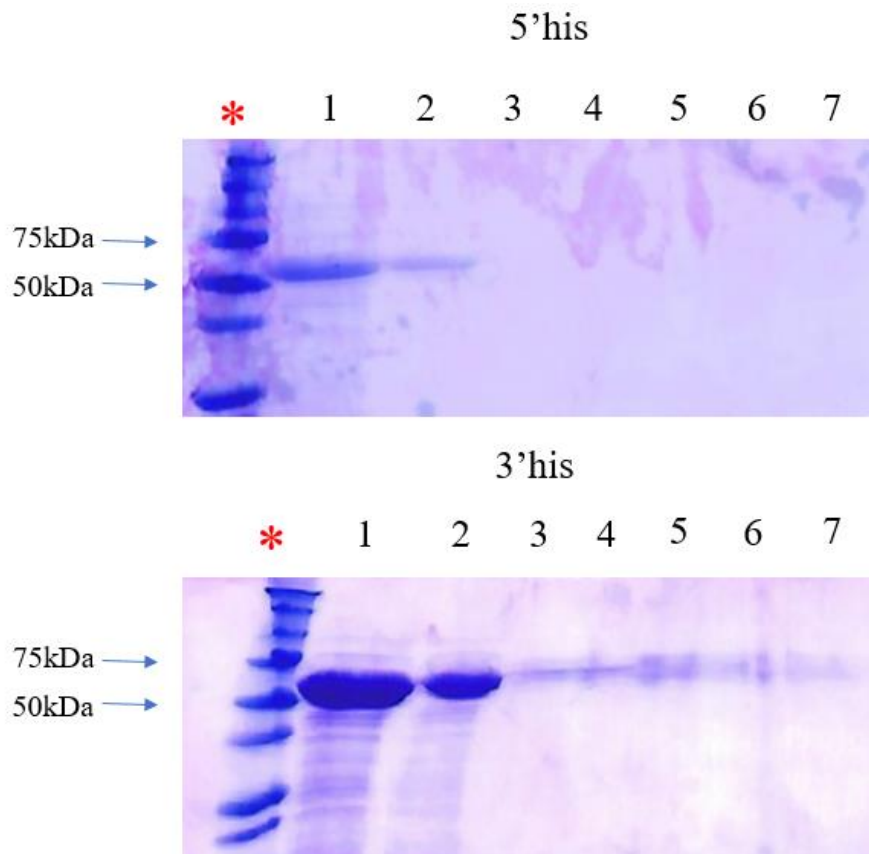
**Figure 9 SDS PAGE Analysis of  $\text{Ni}^{+2}$  based purification of the 5'his his-tagged protein, non-denaturing conditions.** 1) Crude extract 2) Flowthrough 3) Wash 4) 25 mM Imidazole 5) 50mM Imidazole 6) 100 mM Imidazole 7) 200 mM imidazole. Ladder marked with an asterisk.

### 3.1.5.3 Preliminary enzyme assays

Crude extracts of 5'his were assayed for glycosidase activity, and very low activity was observed (data not shown). Moreover, no ginsenosidase activity was detected. This suggested that the addition of the 6× His-tag to the N-terminus may interfere with the active site by steric or electrostatic interactions therefore having an impact on the catalytic properties of the enzyme (Kurtzman, 2009).

### 3.1.5.4 Denatured His-tagged protein extracts

Protein extracts (5'his and 3'his) were prepared using denaturing conditions in an attempt to expose the 6× His Tag, making it available to be bound to the nickel column. As part of the Ni-column purification process, an on-column refolding step was planned. However, despite these precautions, the recombinant proteins of ~66 kDa still eluted in the flow-through and wash fractions (**Figure 10**). Consequently, despite the purification under denaturing conditions using an urea-based buffer and imidazole as the eluent, which can allow for the isolation of biologically active products (Bornhorst and Falke, 2000), the denaturing process of the samples was not successful on properly exposing the 6× His Tag.



**Figure 10 SDS PAGE Analysis of Ni<sup>2+</sup> based purification of the 5'his and His 3 his-tagged proteins, denaturing conditions.** 1) Flowthrough 2) Wash 3) 25 mM Imidazole 4) 50 mM Imidazole 5) 100 mM Imidazole 6) 200 mM imidazole. Ladders marked with an asterisk.



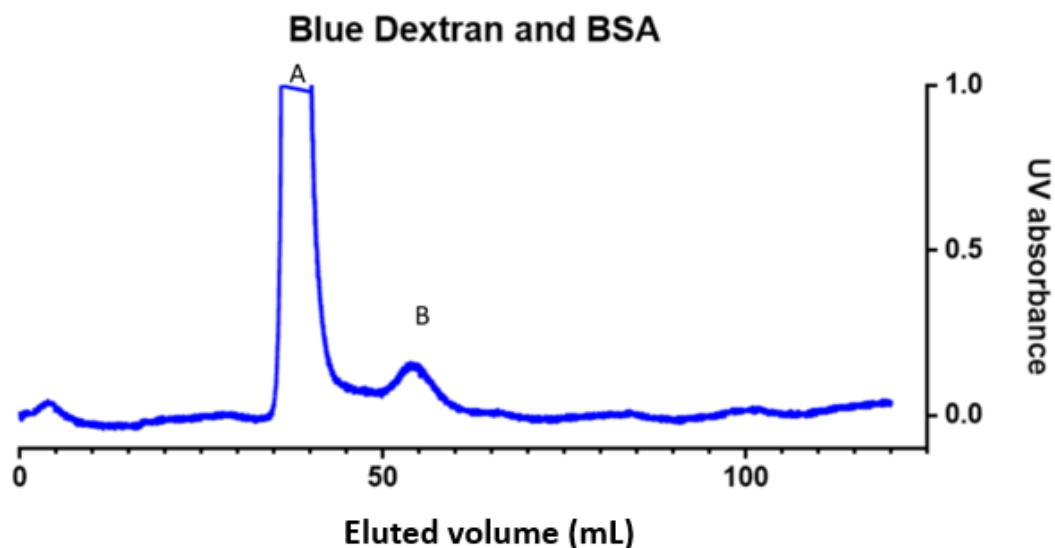
### 3.1.6 Non-tagged proteins

#### 3.1.6.1 Gel Filtration Chromatography

Since the recombinant PiGH1-x is highly expressed, and Ni-affinity column purification was unsuccessful, a partially purified protein preparation was prepared using gel filtration chromatography on a Sephacryl S-200 HR column. This column has a separation range of 5 – 250 kDa, and was expected to separate the recombinant PiGH1-x (~ 66 kDa) from the bulk of the other cellular proteins in the extract.

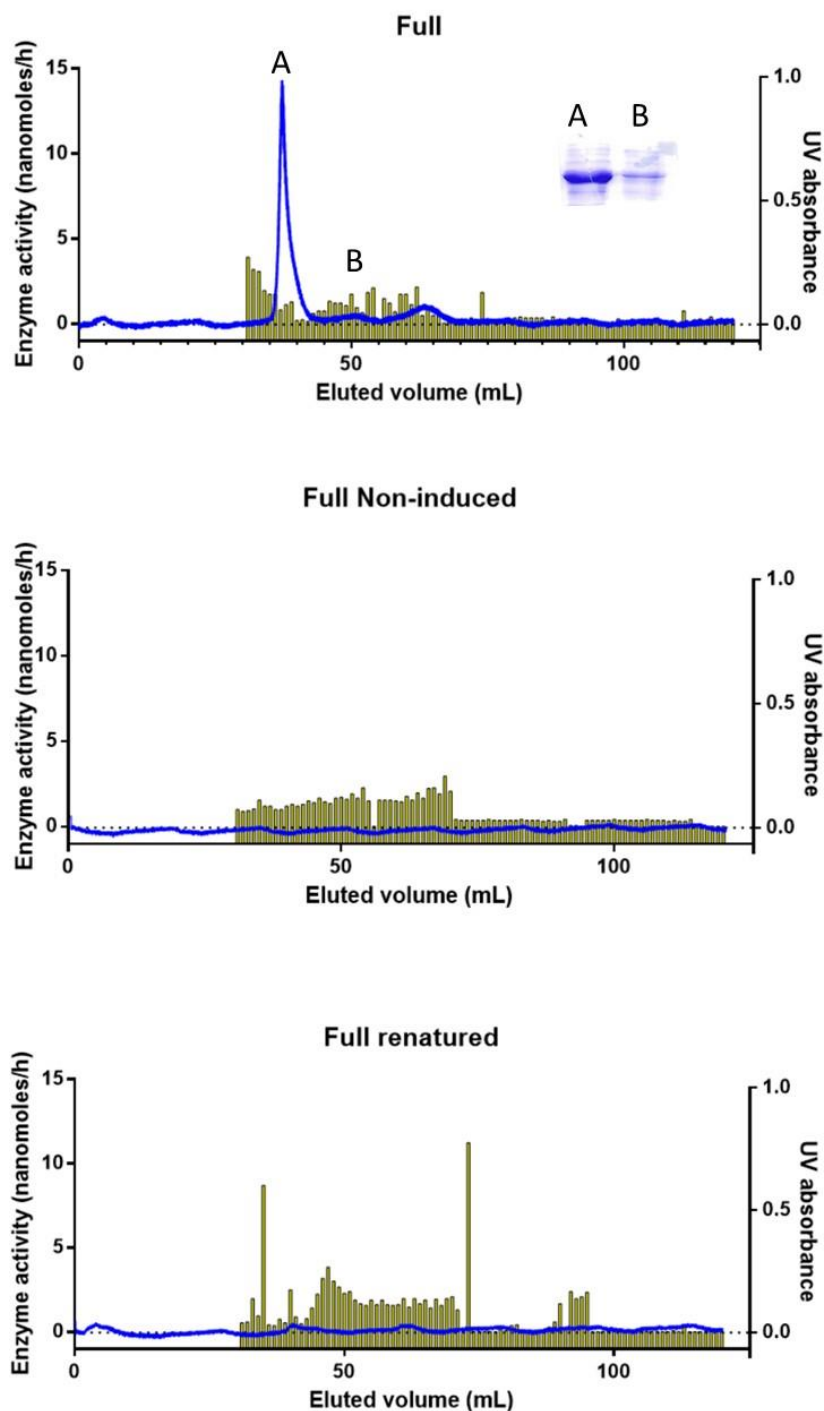
Partial calibration of the Sephacryl S-200 HR column with Blue Dextran and BSA confirmed that the fractions corresponding to the void volume ( $> 2e^6 M_r$ ) eluted between 70-90 min after the beginning of the run (elution volume between 35-45 mL), while fractions corresponding to a size of ~66 kDa after 50-60 mL of elution (**Figure 11**). However, when the protein preparation of the full PiGH1-x recombinant protein was run through the Sephacryl S-200 HR column, the largest peak corresponded to the void volume (35-44 mL), with another two minor peaks detected: one at between 45-55 mL and the other at between 55-70 mL. The peak eluting between 15-25 mL range was expected to contain the monomeric form of the full PiGH1-x protein, which was further verified using 12% SDS-PAGE electrophoresis (**Figure 12**). A protein preparation from a non-induced culture did not yield any relevant UV trace (**Figure 12**).

That the major peak in the UV trace for the Full PiGH1-x recombinant protein run through the Sephacryl S-200 HR column coincided with the void volume was a clear indication that the majority of the recombinant protein was in an aggregated form. Enzyme assays using pNPG as substrate confirmed that both the void volume fractions as well as the fractions corresponding to proteins in the 66 kDa  $M_r$  range contained glycosidase activity. These results suggested that the Full PiGH1-x recombinant protein was forming a large aggregate, and helped explain the lack of His-tag recognition by the anti-His antibody. It also provided guidance for attempting to isolate the Full PiGH1-x recombinant protein under mild denaturing conditions.



**Figure 11 Calibration UV trace using BSA and Blue Dextran.** A) Void volume peak, tracked by addition of Blue dextran. B) BSA peak, corresponding to 66 kDa molecular weight range.

Furthermore, after buffer exchange, the Full Renatured preparation was also submitted to Gel Filtration Chromatography (**Figure 12**). However, the UV trace did not show any significant signal. This outcome suggests a lack of protein in the preparation; which might be explained by the fact that renaturation of the protein through dialysis was incomplete. Indeed, when the dialysate was centrifuged prior to loading on the Sephacryl S-200 HR column, a visible pellet was observed, suggesting that much of the expressed protein was in an insoluble form.



**Figure 12 UV traces and enzyme activity plots of gel filtration chromatography fractions.** Yellow bars correspond to glycosidase activity measured using the pNPG assay, activity is measured in nanomoles of pNP per hour (left Y axis). Blue line corresponds to the UV signal from the Gel Filtration Chromatography (right Y axis). In inset gel in the panel for the full protein shows the presence of a 66 kDa protein in both the void volume fraction (labelled “A”) and 66 kDa fraction (labelled “B”).

### 3.1.6.2 Fraction selection

**Figure 12** shows the activity of each fraction obtained from preparation “full” as well as its respective UV trace. The Full Non-induced preparation was used as a control to measure potential background activity from bacterial glycosidases. Highest activity is observed at the largest fractions, corresponding to the void volume; while fractions on the 66 kDa range also have high glycosidase activity. These fractions were pooled together for subsequent experiments and denominated “66 kDa”, as they comprise the molecular range where recombinant PiGH1-x is expected to elute. However, when comparing to the non-induced control, there was no substantial difference in the pNPG-based activity between the “66 kDa” fractions, while there is a clear increase of activity on the fractions that correspond to the void volume. These data confirmed that recombinant PiGH1-x formed large molecular weight aggregates after expression.

There was a clear activity peak in the 66 kDa range in the Full Renatured preparation, despite the low quantity of protein in the fractions. These resulted from the fact that part of the extract remained insoluble and in the form of a pellet when centrifuged before submission through Gel Filtration, as described above.

## 3.2 Enzymatic activity assays

The nomenclature and the origin of the samples used in the enzymatic activity assays are described in **Table 2**.

**Table 2 Nomenclature and description of protein preparations for enzyme assays**

NAME	DESCRIPTION
Full Void Volume	rPiGH1-x expressed from <i>full</i> construct. Fractions which eluted between 15-44 mL from gel filtration chromatography, corresponding to the void volume.
Full 66 kDa	rPiGH1-x expressed from <i>full</i> construct. Fractions from the 45-55 mL range resulting from gel filtration chromatography.
Full Renatured Void Volume	rPiGH1-x expressed from <i>full</i> construct and extracted under denaturing conditions, renatured by buffer exchange. Fractions which eluted between 15-44 mL from gel filtration chromatography, corresponding to the void volume.
Full Renatured 66 kDa	rPiGH1-x expressed from <i>full</i> construct and extracted under denaturing conditions, renatured by buffer exchange. Fractions from the 45-55 mL range resulting from gel filtration chromatography.
Full Non-induced Void Volume	Protein extract from cultures of cells containing <i>full</i> construct, Non-induced protein expression. Fractions which eluted between 15-44 mL from gel filtration chromatography, corresponding to the void volume.
Full Non-induced 66 kDa	Protein extract from cultures of cells containing <i>full</i> construct, Non-induced protein expression. Fractions from the 45-55 mL range resulting from gel filtration chromatography.
3'his Non-induced	Protein extract from cultures of cells containing 3'his construct, Non-induced protein expression.
3'his Induced	rPiGH1-x expressed from 3'his construct.
3'his Renatured	rPiGH1-x from 3'his construct, recovered from the flowthrough fraction eluted from the nickel affinity column (denaturing conditions). Sample was renatured under the same conditions mentioned in section 2.1.9.
5'his Non-induced	Protein extract from cultures of cells containing 5'his construct, Non-induced protein expression.
5'his Induced	rPiGH1-x expressed from 5'his construct.
5'his Renatured	rPiGH1-x from 5'his construct, recovered from the flowthrough fraction eluted from the nickel affinity column (denaturing conditions). Sample was renatured under the same conditions mentioned in section 2.1.9.

### 3.2.1 Glycosidase Activity Assay

Pooled fractions from the various Sephacryl column runs were tested for glycosidase activity using the pNPG colorimetric assay (**Table 3**). One-way ANOVA (p-value of 0.05) with Sidak's post hoc test were used to assess whether activity detected in induced protein extract samples was significantly different from the equivalent fractions collected for non-induced controls of the same construct. This analysis showed that all induced protein preparations had greater glycosidase activity than those from non-induced cultures, supporting the conclusion that the recombinant PiGH1-x protein expressed in induced cultures is a glycosidase.

The Full Void Volume sample had less activity than the 66 kDa fraction for the same protein preparation, suggesting that the aggregated protein was compromised in terms of catalytic activity.

Furthermore, although protein yields in the Full Renatured preparation were very low, glycosidase activity was observed at the expected molecular weight. And, although activity appears to be higher than in the Non-induced control, this difference is not statistically significant.

In addition, 3'his and 5'his Renatured preparations did not show any glycosidase activity, which means that the denaturing conditions used for their extraction were too harsh and did not allow proper refolding, causing structural changes which inhibited the enzyme activity. On the other hand, 3'his Induced and 5'his Induced, which were prepared from crude extracts did show significant glycosidase activity when compared to their respective Non-induced controls.

**Table 3 Glycosidase activity of protein preparations.** Enzyme specific activity of protein preparations, measured in nmol of PNP/h/ $\mu$ g of protein preparation. Preparation “full void volume non-induced” was used as a control for samples “full void volume” and “full renatured void volume”; “full 66kDa non-induced” was used as a control for “full 66 kDa” and “full renatured 66 kDa”; “3’his non-induced” was used as a control for “3’his induced” and “3’his renatured”; “5’his non-induced” was used as a control for “5’his induced” and “5’his renatured”. Values statistically significant different against their control are presented in bolded numbers, while those which represent a statistically significant increase in activity (relative to their respective control) are marked with an asterisk.

Treatment	Non-induced control	Induced enzyme
Full Void Volume	$0.40 \pm 0.03$	<b><math>0.59 \pm 0.01^*</math></b>
Full Renatured Void Volume	$0.40 \pm 0.03$	$<0.01^*$
Full 66 kDa	$0.10 \pm 0.01$	<b><math>0.98 \pm 0.01^*</math></b>
Full Renatured 66 kDa	$0.10 \pm 0.01$	$0.24 \pm 0.01$
3’his Induced	$0.38 \pm 0.02$	<b><math>0.98 \pm 0.08^*</math></b>
3’his Renatured	$0.38 \pm 0.02$	$0.01 \pm 0.01^*$
5’his Induced	$0.02 \pm 0.01$	<b><math>0.31 \pm 0.10^*</math></b>
5’his Renatured	$0.02 \pm 0.01$	$0.01 \pm 0.01$

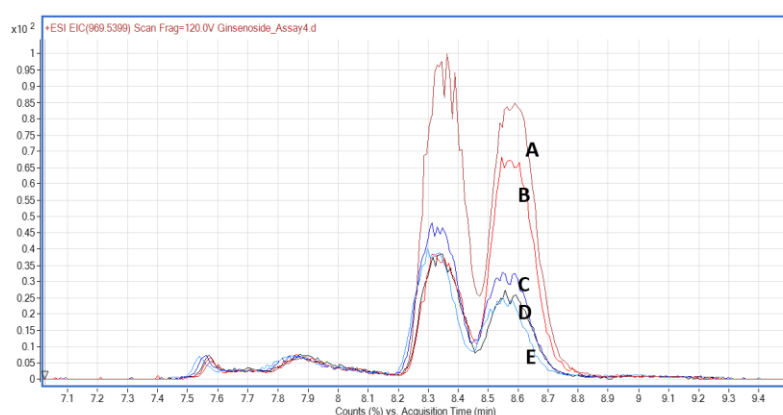
### 3.2.2 Ginsenosidase activity

For ginsenosidase activity measurements, pooled fractions were incubated with a preparation of ginsenoside Rb1, and analysed by LC-MS. Extracted ion chromatograms (EICs) generated for the expected products (969.5399 m/z for Rd and Gypenoside XVII, and 807.4871 m/z for F2) revealed that in some fractions, Rb1 was converted into ginsenosides Rd as well as Gypenoside XVII. One-way ANOVA (p-value of 0.05) with Sidak’s post hoc test were used to evaluate conversion to Rd and conversion to Gypenoside XVII, comparisons were made to assess whether activity detected in induced protein extract samples was significantly different from the equivalent fractions collected for Non-induced controls of the same construct. (**Figure 13, Table 4**). This means that

recombinant PiGH1-x has the ability to partially deglycosylate Rb1, not only via  $\beta(1\rightarrow2)$  glycosidase activity, as initially proposed (Ivanov, 2015), but also via  $\beta(1\rightarrow6)$  glycosidase activity (**Table 5**). However, there was no F2 formation, which would have indicated both  $\beta(1\rightarrow2)$  and  $\beta(1\rightarrow6)$  glycosidase activity acting on the same Rb1 substrate.

The fact that significant ginsenosidase activity was detected in the Full Void Volume sample, confirmed the initial conclusion that the recombinant protein was aggregating and eluting as a high molecular fraction. Moreover, since higher activity was measured in the Full Renatured Sample, it can be concluded that the aggregated protein does not have as much activity as it would in its monomeric form.

Although the pNPG-based activity of the Full Renatured 66 kDa sample was not statistically significant from other glycosidase preparations (**Table 3**), this preparation had significantly higher ginsenosidase activity than in the corresponding non-induced preparation (**Table 4**). In the case of the Full Renatured 66 kDa sample, the background protein composition may have been modified. Moreover, a part of the sample was lost during the concentration process. Since the ginsenosidase activity of the Full Renatured 66 kDa sample is statistically significant, this could indicate that the glycosidase activity observed comes mainly from rPiGH1-x.



**Figure 13 Raw EIC traces.** EIC (Extracted Ion Chromatograms) @ 969.5399 m/z. LC-MS detection of ionic abundance of reaction products. First peak (8.4 min) corresponds to Rd. Second peak (8.6 min) corresponds to GXVII. Black (D) Mock assay. Dark Blue (C) Full Non-induced Void Volume. Light Blue (E) Full Non-induced 66kDa. Dark Red (A) Full induced Void Volume. Light Red (B) Full induced 66kDa.



**Table 4 Ginsenosidase activity of protein preparations.** Measured in nmol of product/h/mg of protein preparation. Preparation Full Void Volume Non-induced was used as a control for samples Full Void Volume and Full Renatured Void Volume; Full 66kDa Non-induced was used as a control for Full 66 kDa and Full Renatured 66 kDa; 3'his Non-induced was used as a control for 3'his Induced and 3'his Renatured; 5'his Non-induced was used as a control for 5'his Induced and 5'his Renatured. Values statistically significant different against their control are presented in bolded numbers, while those which represent a statistically significant increase in activity are marked with an asterisk.

Reaction Product	Treatment	Non-induced control	Induced Enzyme
<b>Rd</b>	Full Void Volume	3.92 ± 0.15	<b>17.23 ± 0.40*</b>
	Full Renatured Void Volume	3.92 ± 0.15	4.10 ± 0.36
	Full 66 kDa	1.52 ± 0.02	1.01 ± 0.04
	Full Renatured 66 kDa	1.52 ± 0.02	<b>37.99 ± 3.45*</b>
	3'his Induced	0.75 ± 0.01	<b>10.13 ± 1.42*</b>
	3'his Renatured	0.75 ± 0.01	0.21 ± 0.01
	5'his Induced	<0.01	<0.01
	5'his Renatured	<0.01	<0.01
<b>GXVII</b>	Full Void Volume	3.15 ± 0.21	<b>19.11 ± 0.12*</b>
	Full Renatured Void Volume	3.15 ± 0.21	7.55 ± 0.46
	Full 66 kDa	<0.01	<b>13.06 ± 0.18*</b>
	Full Renatured 66 kDa	<0.01	<b>30.54 ± 3.12*</b>
	3'his Induced	4.63 ± 0.46	<b>13.17 ± 0.95*</b>
	3'his Renatured	4.63 ± 0.46	2.34 ± 0.05
	5'his Induced	1.16 ± 0.03	1.84 ± 0.12
	5'his Renatured	1.16 ± 0.03	1.22 ± 0.09

**Table 5 Summary of resulting Glycosidase activity based on substrate conversion.**  
 Conversion of Ginsenoside Rb1 to Rd demonstrates (1→2) glycosidase activity, while conversion of Ginsenoside Rb1 to GXVII represents (1→6) activity.

<b>Treatment</b>	<b>Glycosidase activity</b>
Full Void Volume	(1→2) and (1→6)
Full Renatured Void Volume	No
Full 66 kDa	(1→2)
Full Renatured 66 kDa	(1→2) and (1→6),
3'his Induced	(1→2) and (1→6),
3'his Renatured	No
5'his Induced	No
5'his Renatured	No

## 4 GENERAL DISCUSSION

The objective of this study was to clone and express the recombinant protein PiGH1-x, from the sequence obtained by Ivanov (2015). Furthermore, this work aimed to test the ability of the PiGH1-x to deglycosylate ginsenosides. The protein expressed proved to be of the expected size of around 66.1 kDa (Ivanov, 2015).

Protein expression was achieved using *E.coli* DH5 $\alpha$  transformed with the expression plasmid pET28a containing the PiGH1-x ORF. Using gel filtration chromatography, it was demonstrated that most of the protein formed aggregates during the expression process. This issue may be due to the fact that bacteria do not have the capacity to perform post-translational modifications by chaperones and proteases, which in many cases are needed to achieve ideal eukaryotic protein function (Allen et al., 1992). Additionally, high-level protein production and therefore accumulation, can result in the formation of intracellular electron-dense protein aggregates best known as inclusion bodies (Allen et al., 1992). These bacterial inclusion bodies are the result of an imbalance between protein deposition and release, usually generated by the failure in folding or degradation of misfolded polypeptides (Villaverde and Carrió, 2003). This phenomenon may be avoided by heterologous expression in a eukaryotic organism, such as *Komagataella phaffii*, which is a well characterised organism for heterologous expression of recombinant proteins (Ahmad et al., 2014) (previously known as *Pichia pastoris* (Kurtzman, 2009)).

Even though the aggregated protein contained in the “void volume fractions” (“full” and “full renatured”) had reduced functionality, these fractions could deglycosylate ginsenoside Rb1, converting it into Rd as well as into Gypenoside XVII, suggesting that there are both  $\beta(1\rightarrow2)$  and  $\beta(1\rightarrow6)$  glycosidase activities in these fractions. The same pattern of ginsenosidase activity, albeit at much lower levels, was observed for both of the 5’his and 3’his constructs in crude extracts.

In the case of the Gel Filtration Chromatography fractions corresponding to 66 kDa proteins, the “Full” protein generated Gypenoside XVII, demonstrating only  $\beta(1\rightarrow2)$  glycosidase activity as initially hypothesized (Ivanov, 2015). In contrast, when subjected

to the denaturing/renaturing, process the “full renatured” sample demonstrated both  $\beta(1\rightarrow2)$  and  $\beta(1\rightarrow6)$  glycosidase activities. It is possible that the denaturing/renaturing process modified the protein active site, conferring  $\beta(1\rightarrow6)$  glycosidase activity.

The fact that recombinant PiGH1-x has ginsenosidase activity supports the hypothesis that this enzyme has a role in the selective deglycosylation of the 20(*S*)-protopanaxadiol ginsenoside Rb1 by *P. irregulare* (Yousef and Bernards, 2006). Most likely, the organism uses PiGH1-x to partially degrade ginsenosides, which the ginseng plant uses as a barrier to prevent infection by other pathogens (Bowyer et al., 1995). Furthermore, ginsenosides can promote the growth of *P. irregulare* (Yousef and Bernards, 2006), which may happen via the partial metabolism of ginsenosides. If this occurs in the field, it may allow *P. irregulare* to accumulate inoculum and therefore facilitate the infection process in subsequent ginseng crops. Additionally, the fact that the protein extracts used as the starting material in the prediction of the sequence of PiGH1-x were expressed by *P. irregulare* when exposed to ginsenosides (Neculai et al., 2009), further supports the idea that this enzyme has a role in the pathogenicity of the microorganism against ginseng.

## 5 CONCLUSION AND FUTURE RECOMMENDATIONS

A recombinant PiGH1-x protein was highly expressed in a heterologous (*E. coli*) system, although the process still needs optimization. Specifically, the production of high levels of recombinant PiGH1-x in *E. coli* DH5 $\alpha$  resulted in protein aggregates, possibly due to a lack of post-translational modifications. Therefore, recombinant expression in a different organism with the capability to produce a correctly folded protein, is needed in order to obtain a recombinant protein with the correct structure.

Nonetheless, even though the ideal protein conformation was not achieved, the  $\beta(1\rightarrow2)$  ginsenosidase activity of recombinant PiGH1-x was demonstrated, as it retained part of its catalytic properties. Remarkably, when recombinant PiGH1-x was extracted under denaturing conditions, and subsequently renatured, both  $\beta(1\rightarrow2)$  and  $\beta(1\rightarrow6)$  ginsenosidase activities were observed. Further enzyme assays using a “correctly” folded recombinant protein would be needed to confirm its glycosidase specificity ( $\beta(1\rightarrow2)$ ,  $\beta(1\rightarrow6)$  or both), which will allow the substitution of the place-holder “x” in the name of the protein with the number that corresponds to its specific catalytic activity.

Confirmation that the protein sequence obtained by (Neculai et al., 2009) corresponds to a ginsenosidase, as well as the steps taken into the optimization of its expression as a recombinant protein, are important steps in elucidating the role of extracellular glycosidases of *P. irregulare* in its pathogenicity towards ginseng. Optimization of both the protein induction and its subsequent purification are required to obtain sufficient protein for full characterization.

Enzyme kinetic parameters (i.e.  $V_{\max}$  and  $K_M$ ), should be determined to further characterise the enzyme. Moreover, additional enzyme assays should be performed using different substrates (e.g. other saponins) to evaluate the ability of PiGH1-x to deglycosylate other molecules. This knowledge would provide a better idea of how this protein works, therefore allowing to better elucidate its mechanism of action.

Additionally, it would be a key step to measure the induced expression of the protein, at the gene level, in response to treatment of *P.irregulare* culture with ginsenosides. This would confirm that PiGH1-x is expressed by the organism when exposed to ginseng, as part of its infection mechanism.

Knowledge of the specific mechanisms that confer the pathogenicity of an organism makes it more probable to succeed in the suppression of diseases, as it allows a better design of controlling agents and processes (Martin and Loper, 1999). Therefore, in order to solve some of the current agronomical issues related to the production of American ginseng, such as the re-plant problem, it is of great importance to continue studying all of the pieces that play a role on its infection by *P. irregulare*.

## 6 REFERENCES

- Adams, P. B. (1971). *Pythium aphanidermatum* oospore germination as affected by time, temperature, and pH. *Phytopathology*, 61(9), 1149–1150.
- Agrios, G. N. (2005). *Plant Pathology* (4th ed.). San Diego: Academic Press.
- Ahmad, M., Hirz, M., Pichler, H., and Schwab, H. (2014). Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Applied Microbiology and Biotechnology*, 98(12), 5301.
- Allen, S. P., Polazzi, J. O., Gierse, J. K., and Easton, A. M. (1992). Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*. *Journal of Bacteriology*, 174(21), 6938–6947.
- Ang-lee, M. K., Moss, J., and Yuan, C.-S. (2001). Herbal medicines and perioperative care. *Journal of the American Medical Association*, 286(2), 208–216.
- Baldauf, S. L., Roger, A. J., Wenk-Siefert, I., and Doolittle, W. F. (2000). A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science (New York, N.Y.)*, 290(5493), 972–977.
- Bernards, M. A., Ivanov, D. A., Neculai, A. M., and Nicol, R. W. (2011). Ginsenosides: Phytoanticipins or Host Recognition Factors? In D. R. Gang (Ed.), *The Biological Activity of Phytochemicals* (pp. 13–30). Pullman: Springer.
- Birch, P. R. ., Rehmany, A. ., Pritchard, L., Kamoun, S., and Beynon, J. (2006). Trafficking arms: oomycete effectors enter host plant cells. *Trends in Microbiology*, 14, 8–11.
- Bornhorst, J., and Falke, J. (2000). Purification of proteins using polyhistidine affinity tags. *Methods in Enzymology*, 326, 245–254.
- Bowyer, P., Clarke, B. R., Lunnes, P., Daniels, M. J., and Osbourn, A. E. (1995). Host range of a plant pathogenic fungus determined by a saponin detoxifying enzyme.

- American Association for the Advancement of Science*, 267(5196), 371–374.
- Brekhman, I. I., and Dardymov, I. V. (1969). New substances of plant origin which increase nonspecific resistance. *Annual Review of Pharmacology*, 9, 419–430.
- Cheng, L. Q., Ju-Ryun, N., Kim, M. K., Bang, M. H., and Yang, D. C. (2007). Microbial conversion of ginsenoside Rb1 to minor ginsenoside F2 and gypenoside XVII by *Intrasporangium* sp. GS603 isolated from soil. *Journal of Microbiology and Biotechnology*, 17(12), 1937–1943.
- Cui, S., Wu, J., Wang, J., and Wang, X. (2017). Discrimination of American ginseng and Asian ginseng using electronic nose and gas chromatography-mass spectrometry coupled with chemometrics. *Journal of Ginseng Research*, 41(1), 85–95.
- Dean, R. A., Lichens-Park, A., and Kole, C. (2014). *Genomics of Plant-Associated Fungi and Oomycetes: Dicot Pathogens*. Heidelberg: Springer. Retrieved from
- Debeljak, N., Feldman, L., Davis, K. L., Komel, R., and Sytkowski, A. J. (2006). Variability in the immunodetection of His-tagged recombinant proteins. *Analytical Biochemistry*, 359(2), 216–223.
- Fuzzati, N., Gabetta, B., Jayakar, K., Pace, R., and Peterlongo, F. (1999). Liquid chromatography – electrospray mass spectrometric identification of ginsenosides in *Panax ginseng* roots. *Journal of Chromatography A*, 854, 69–79.
- Günther, S., Varma, A. K., Moza, B., Kasper, K. J., Wyatt, A. W., Zhu, P., ... Sundberg, E. J. (2007). A novel loop domain in superantigens extends their T cell receptor recognition site. *Journal of Molecular Biology*, 371(1), 210–221.
- Hancock, J. G. (1977). Factors affecting soil populations of *Pythium ultimum* in the San Joaquin Valley of California. *Hilgardia*, 45(4), 107–122.
- Hardham, A. and Hyde, G. (1997). Asexual sporulation in the Oomycetes. *Advances in Botanical Research*, 24, 353–398.



- Hardham, A. R. (2007). Cell biology of plant-oomycete interactions. *Cellular Microbiology*, 9(1), 31–39.
- Harvey, P. R., Butterworth, P. J., Hawke, B. G., and Pankhurst, C. E. (2001). Genetic and pathogenic variation among cereal, medic and sub-clover isolates of *Pythium irregulare*. *Mycological Research*, 105(1), 85–93.
- Hendrix, F. F. and Campbell, W. A. (1973). *Pythiums* as plant pathogens. *Annual Review of Phytopathology*, 11(1), 77–98.
- Howard, R., Garland, J., and Seaman, W. (1994). *Diseases and Pests of Vegetable Crops in Canada*. Ottawa: Canadian Phytopathological Society and Entomological Society of Canada.
- Isaac, S. (1992). *Fungal-plant interactions*. London: Chapman and Hall.
- Ivanov, D.A. (2015) Ginsenosides, Glycosidases and the Ginseng-Pythium Interaction, School of Graduate and Postdoctoral Studies, University of Western Ontario.
- Ivanov, D.A. and Bernards, M.A. (2012) Ginsenosides and the pathogenicity of *Pythium irregulare*. *Phytochemistry*, 78, 44-53.
- Ivanov, D. A., Georgakopoulos, J. R. C., and Bernards, M. A. (2015). The chemoattractant potential of ginsenosides in the ginseng - *Pythium irregulare* pathosystem. *Phytochemistry*, 122, 56–64.
- Johnson, L. F. (1988). Effects of atmospheric gases and light on changes in thickness of oospore walls and on germinability of oospores of *pythium ultimum*. *Ecology and Epidemiology*, 78(4), 435–439.
- Johnson, L. F., and Arroyo, T. (1983). Germination of oospores of *Pythium ultimum* in the cotton rhizosphere. *Phytopathology*, 73(12), 1620-1624.
- Kernaghan, G., Reeleder, R. D., and Hoke, S. M. T. (2007). Quantification of *Cylindrocarpon destructans* f. sp. panacis in soils by real-time PCR. *Plant*

- Pathology*, 56(3), 508–516.
- King, M. L., Adler, S. R., and Murphy, L. L. (2006). Extraction-dependent effects of American ginseng (*Panax quinquefolium*) on human breast cancer cell proliferation and estrogen receptor activation. *Integrative Cancer Therapies*, 5(3), 236–43.
- Kurtzman, C. P. (2009). Biotechnological strains of *Komagataella* (*Pichia*) *pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis. *Journal of Industrial Microbiology and Biotechnology*, 36(11), 1435–1438.
- Larkin, R. P., English, J. T., and Mihail, J. D. (1995). Identification, distribution and comparative pathogenicity of *Pythium* spp associated with alfalfa seedlings. *Soil Biology and Biochemistry*, 27(3), 357–364.
- Latijnhouwers, M., De Wit, P. J. G. M., and Govers, F. (2003). Oomycetes and fungi: Similar weaponry to attack plants. *Trends in Microbiology*, 11(10), 462–469.
- Lifshitz, R., and Hancock, J. G. (1984). Environmental influences on the passive survival of *Pythium ultimum* in soil. *Phytopathology*, 76 (2), 128-132.
- Lifshitz, R., Windam, M. T., and Baker, R. (1986). Mechanism of biological control of preemergence damping-off of pea by seed treatment with *Trichoderma* spp. *Phytopathology*, 76 (7), 720-725.
- Lumsden, R. D., and Ayers, W. A. (1975). Influence of soil environment on the germinability of constitutively dormant oospores of *Pythium ultimum*. *Phytopathology*, 65(10), 1101–1107.
- Martin, F. N., and Loper, J. E. (1999). Soilborne Plant Diseases Caused by *Pythium* spp.: ecology, Epidemiology, and Prospects for Biological Control. *Critical Reviews in Plant Sciences*, 18(2), 111–181.
- Matsumoto, C., Kageyama, K., Suga, H., and Hyakumachi, M. (2000). Intraspecific DNA polymorphisms of *Pythium irregulare*. *Mycological Research*, 104(11), 1333–1341.

- McCormick, J. K., and Schlievert, P. M. (2003). Expression, Purification and Detection of Novel Streptococcal Superantigens. In T. Krakauer (Ed.), *Methods in Molecular Biology*, 214, 33–45). Totowa, NJ: Humana Press Inc.
- Neculai, A. M., Ivanov, D., and Bernards, M. A. (2009). Partial purification and characterization of three ginsenoside-metabolizing betha-glucosidases from *Pythium irregulare*. *Phytochemistry*, 70 , 1948–1957.
- Nelson, E. B. (1990). Exudate molecules initiating fungal responses to seeds and roots. *Plant and Soil*, 129, 197–209.
- Nelson, E. B., and Craft, C. M. (1989). Comparative germination of culture-produced and plant-produced sporangia of *Pythium ultimum* in response to soluble seed exudates and exudate components. *Phytopathology*, 79(10), 1009-1013.
- Nicol, R. W., Traquair, J. A., and Bernards, M. A. (2002). Ginsenosides as host resistance factors in American ginseng ( *Panax quinquefolius* ). *Canadian Journal of Botany*, 80, 557–562.
- Nicol, R. W., Yousef, L., Traquair, J. A., and Bernards, M. A. (2003). Ginsenosides stimulate the growth of soilborne pathogens of American ginseng. *Phytochemistry*, 64(1), 257–264.
- Qi, L., Wang, C., and Yuan, C. (2011). Ginsenosides from American ginseng : chemical and pharmacological diversity. *Phytochemistry*, 72(8), 689–699.
- Quidde, T., Büttner, P., and Tudzynski, P. (1999). Evidence for three different specific saponin-detoxifying activities in *Botrytis cinerea* and cloning and functional analysis of a gene coding for a putative avenacinase. *European Journal of Plant Pathology*, 105, 273–283.
- Quidde, T., Osbourn, A. ., and Tudzynski, P. (1998). Detoxification of  $\alpha$ -tomatine by *Botrytis cinerea*. *Physiological and Molecular Plant Pathology*, 52, 151–165.
- Rahman, A. K. M., Bonsor, D. A., Herfst, C. A., Pollard, F., Peirce, M., Wyatt, A. W.,

- Kasper, K.J., Madrenas, J., Sundberg, E.J., McCormick, J. K. (2011). The T cell receptor beta-chain second complementarity determining region loop (CDR2beta) governs T cell activation and Vbeta specificity by bacterial superantigens. *Journal of Biological Chemistry*, 286(6), 4871–4881.
- Roy, R., Grohs, R., and Reeleder, R. (2003). A method for the classification by shape of dried roots of ginseng (*Panax quinquefolius* L.). *Canadian Journal of Plant Science*, 83(4), 955–958.
- Schlag, E. M., and McIntosh, M. S. (2006). Ginsenoside content and variation among and within American ginseng (*Panax quinquefolius* L.) populations. *Phytochemistry*, 67(14), 1510–1519.
- Teng, R., Li, H., Chen, J., Wang, D., He, Y., and Yang, C. (2002). Complete assignment of <sup>1</sup>H and <sup>13</sup>C NMR data for nine protopanaxatriol glycosides. *Magnetic Resonance in Chemistry*, 40, 483–488.
- Villaverde, A., and Carrió, M. M. (2003). Protein aggregation in recombinant bacteria: Biological role of inclusion bodies. *Biotechnology Letters*, 25(17), 1385–1395.
- Wen, J., and Zimmer, E. a. (1996). Phylogeny and biogeography of *Panax* L. (the ginseng genus, araliaceae): inferences from ITS sequences of nuclear ribosomal DNA. *Molecular Phylogenetics and Evolution*, 6(2), 167–177.
- Yang, M., Zhang, X., Xu, Y., Mei, X., Jiang, B., Liao, J., Yin, Z., Zheng, Z., Zhao, Z., Fan, L., He, X., Zhu, S. (2015). Autotoxic ginsenosides in the rhizosphere contribute to the replant failure of *Panax notoginseng*. *PLoS ONE*, 10(2), 1–17.
- Yousef, L. F., and Bernards, M. A. (2006). In vitro metabolism of ginsenosides by the ginseng root pathogen *Pythium irregulare*. *Phytochemistry*, 67(16), 1740–1749.
- Yuk, J., McIntyre, K. L., Fischer, C., Hicks, J., Colson, K. L., Lui, E., Brown, D., Arnason, J. T. (2013). Distinguishing Ontario ginseng landraces and ginseng species using NMR-based metabolomics. *Analytical and Bioanalytical Chemistry*, 405(13), 4499–4509.

## 7 APPENDICES

### Appendix 1 List of primers

<i>full</i> primers	
Forward	<p>5'-TTT<u>CCATGG</u>ATGAAGTGTCTGCAACTG-3'</p> <p><u>NcoI</u>                      <b>PiGH1-x ORF</b></p>
Reverse	<p>5'-CCC<u>GGATCC</u><b>TTACTGTCCACGCACGAG</b>-3'</p> <p><u>BamHI</u>    *                      <b><u>PiGH1-x ORF</u></b></p>
<i>3'his</i> primers	
Forward	<p>5'-TTT<u>CCATGG</u>ATGAAGTGTCTGCAACTG-3'</p> <p><u>NcoI</u>                      <b>PiGH1-x ORF</b></p>
Reverse	<p>5'- CCC<u>GGATCC</u><b>TTAGCCGCTGCTGGATGATGATGATGGCTGCTGC</b><u>CTGTCCACGCACGAG</u>- 3'</p> <p><u>BamHI</u>    *                      6× <i>His</i>                      <b><u>PiGH1-x ORF</u></b></p>

<i>5'his primers</i>			
Forward			
5'-			
TTTCCATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCGAAAACTTG			
TATTTCCAAAGTATGAAGTGTCTGCAACTG-3'			
<u>NcoI</u>	6× His	Tev Site	<u>PiGH1-x</u>
<u>ORF</u>			
Reverse			
5'-CCCGGATCCTTACTGTCCACGCACGAG-3'			
<u>BamHI</u>	*	<u>PiGH1-x ORF</u>	
<i>short primers</i>			
Forward			
5'-			
TTTCCATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCGAAAACTTG			
TATTTCCAAAGTAGCGCTGTCGTCACCA-3'			
<u>NcoI</u>	6× His	Tev Site	<u>tPiGH1-x</u>
<u>ORF</u>			
Reverse			
5'-CCCGGATCCTTACTGTCCACGCACGAG-3'			
<u>BamHI</u>	*	<u>PiGH1-x ORF</u>	

## Appendix 2 Sequence alignment of *full*

Fast alignment of DNA sequences pigh1x.seq and full.seq

Ktuple=2 Gap\_penalty=7

Upper line: pigh1x.seq, from 1 to 1769

Lower line: full.seq, from 10 to 1778

pigh1x.seq:full.seq identity= 100.00%(1770/1770) gap=0.51%(9/1779)

```

1      .....ATGAAGTGTCTGCAACTGCTTCTGGCTGGCGTCTCGGCCGTGCTCGCGACT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1      TTTCCATGGATGAAGTGTCTGCAACTGCTTCTGGCTGGCGTCTCGGCCGTGCTCGCGACT

52     GTGAGTGCAGCGCTGTCGTCACCACCTCCGCGAGTGCCAAGCGCTGCTTCCCGGACGAC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61     GTGAGTGCAGCGCTGTCGTCACCACCTCCGCGAGTGCCAAGCGCTGCTTCCCGGACGAC

112    TTCCTCTTTGGCTCTGCCACCGCCGCATACCAAGTGGAGGGCGCTTGGAACGAGACCGGC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121    TTCCTCTTTGGCTCTGCCACCGCCGCATACCAAGTGGAGGGCGCTTGGAACGAGACCGGC

172    CGAACGCCGTCAATCTGGGACGACTTCTGTGCGGAGAAGCCCGGTCTTCAGTGCTCGAAC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181    CGAACGCCGTCAATCTGGGACGACTTCTGTGCGGAGAAGCCCGGTCTTCAGTGCTCGAAC

232    GTCGCCGACGACTTCTACCACCGGTACGTGAGCGACATCCAGACGATGGTCGACTCGGGA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241    GTCGCCGACGACTTCTACCACCGGTACGTGAGCGACATCCAGACGATGGTCGACTCGGGA

292    TTGGA CTGTTCCGCTTCAGCATCTCCTGGTCGCGCGCCATGAACTGGGACCCTGCGACC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301    TTGGA CTGTTCCGCTTCAGCATCTCCTGGTCGCGCGCCATGAACTGGGACCCTGCGACC

352    AAGAAGATGAAGCCCAACCCGCAAGGCATCGCCTTCTACCACGCGCTCATCGACGAGCTC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361    AAGAAGATGAAGCCCAACCCGCAAGGCATCGCCTTCTACCACGCGCTCATCGACGAGCTC

412    AAGCGCAACAATATCGTGCCGATCTTGACGCTCTACCACTGGGACCTCCCGAGCGAATTG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421    AAGCGCAACAATATCGTGCCGATCTTGACGCTCTACCACTGGGACCTCCCGAGCGAATTG

472    CACAGTGAGCTCAAGCCGCAGGGTTGGCTCAACCGCGAGATCATCGATCACTTCGTTGAG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481    CACAGTGAGCTCAAGCCGCAGGGTTGGCTCAACCGCGAGATCATCGATCACTTCGTTGAG

532    TACTCGGAGCTCATATTCAACGAGTTTGGCAAGAAGGTGGACATGTGGACGACGTTCAAT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541    TACTCGGAGCTCATATTCAACGAGTTTGGCAAGAAGGTGGACATGTGGACGACGTTCAAT

592    GAGCCGCTGTCGTTTCGTCTCCAGGGCTACGCAACGGGCATGTCGGCTCCAGGATACACC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601    GAGCCGCTGTCGTTTCGTCTCCAGGGCTACGCAACGGGCATGTCGGCTCCAGGATACACC

```

652 AACTCCAAGACGAACGCGTACGTTGTCGCGCACAACTGCTGCTCTCGCACGGCAAGGCG  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 661 AACTCCAAGACGAACGCGTACGTTGTCGCGCACAACTGCTGCTCTCGCACGGCAAGGCG  
  
 712 GTGCAGCGCTTCCGCGAGCTCAAGAAGAGCAACGTCGTCCATGAGAAGGCGCGCATCGGT  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 721 GTGCAGCGCTTCCGCGAGCTCAAGAAGAGCAACGTCGTCCATGAGAAGGCGCGCATCGGT  
  
 772 ATCGTGATTGTCGACCACTTCTACTACCCTCTCGACCCGAGCAACCCGCGCGACGTCATT  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 781 ATCGTGATTGTCGACCACTTCTACTACCCTCTCGACCCGAGCAACCCGCGCGACGTCATT  
  
 832 GCGGCTGAGCGCGCGATCCAGTTTGGATTGCGGTGGTTCTTCATTCCCATAGTGACTGGC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 841 GCGGCTGAGCGCGCGATCCAGTTTGGATTGCGGTGGTTCTTCATTCCCATAGTGACTGGC  
  
 892 AAGTATCCAGACGTGATGCGCGAACGTGTCGGTGACCGCCTGCCAGAGTTCACGGCGGAG  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 901 AAGTATCCAGACGTGATGCGCGAACGTGTCGGTGACCGCCTGCCAGAGTTCACGGCGGAG  
  
 952 GAAGCTGCGATGGTAAAGGACTCGTACGACCTCTTCATGCTCAACCACTACTCATCCAC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 961 GAAGCTGCGATGGTAAAGGACTCGTACGACCTCTTCATGCTCAACCACTACTCATCCAC  
  
 1012 ACTGTGACGGACTGCGACTCCGAAACGTGCTCGTCAAGTGCAGCGATTTGTCCCTCGGA  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1021 ACTGTGACGGACTGCGACTCCGAAACGTGCTCGTCAAGTGCAGCGATTTGTCCCTCGGA  
  
 1072 TGGGAACAGGATATGGGCGTCGACAGTACGCGTGCGCCAGACGGTGCGCGTCTGTCCAGC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1081 TGGGAACAGGATATGGGCGTCGACAGTACGCGTGCGCCAGACGGTGCGCGTCTGTCCAGC  
  
 1132 AAGAACAGCCTCGGCGAGTACAACGTGCAAGTGGTTACAGGGATACCCGGCAGGCTACCTG  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1141 AAGAACAGCCTCGGCGAGTACAACGTGCAAGTGGTTACAGGGATACCCGGCAGGCTACCTG  
  
 1192 GAAATGATCAAGTGGATGCACAAGTACAACACAAAGGCGGACATCTTACTCACCGAGAAC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1201 GAAATGATCAAGTGGATGCACAAGTACAACACAAAGGCGGACATCTTACTCACCGAGAAC  
  
 1252 GGCTGGTGCGGCAACGAGACCATCGACAACCAGGACCAACTCTGGTACTTCCAGAGCTAC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1261 GGCTGGTGCGGCAACGAGACCATCGACAACCAGGACCAACTCTGGTACTTCCAGAGCTAC  
  
 1312 ATCGGCGAAGTCTACAAGGCGATCAACGAACTCAAGATCCCCATCATTGGGTACTGCG  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1321 ATCGGCGAAGTCTACAAGGCGATCAACGAACTCAAGATCCCCATCATTGGGTACTGCG  
  
 1372 TGGTCGTTTCATGGACAACCTACGAGTGGGGCTCATTCAAGCCTCGCTTTGGACTCTACTAC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1381 TGGTCGTTTCATGGACAACCTACGAGTGGGGCTCATTCAAGCCTCGCTTTGGACTCTACTAC  
  
 1432 GTCAATTTTCACGTCCGAGACCGGCAAGGCGGATGAGGTACCCCCAGGATCAAACGCCTTG  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1441 GTCAATTTTCACGTCCGAGACCGGCAAGGCGGATGAGGTACCCCCAGGATCAAACGCCTTG  
  
 1492 GCACGCATCCCACGTCTCTGCGGCCAAATGGTACGCGCAAGTGGCGAAGACCAAGTGCCTC



```

|||||
1501 GCACGCATCCCACGTCCTGCGGCCAAATGGTACGCGCAAGTGGCGAAGACCAAGTGCCTC
|||||
1552 GACGGCTTTGAGATTGAAGGCGTGCCAGTCTTGGCAGGAGTCAGCGACGTTACTGGTACC
|||||
1561 GACGGCTTTGAGATTGAAGGCGTGCCAGTCTTGGCAGGAGTCAGCGACGTTACTGGTACC
|||||
1612 AGCTCTGGCAGCAGCTTTAAGACGTTTCGCACTCTGGGGGCTGGTGCTGGGTGTTGTTGTC
|||||
1621 AGCTCTGGCAGCAGCTTTAAGACGTTTCGCACTCTGGGGGCTGGTGCTGGGTGTTGTTGTC
|||||
1672 GCGGTACCAGCGATCGCTTTCTTTGTCAAGAGCCGTCGCTCTCGTCAGAGCTCGCTGCTC
|||||
1681 GCGGTACCAGCGATCGCTTTCTTTGTCAAGAGCCGTCGCTCTCGTCAGAGCTCGCTGCTC
|||||
1732 CGCGCGGGCGAGAACACGCCGCTCGTGCGTGGACAGTAA
|||||
1741 CGCGCGGGCGAGAACACGCCGCTCGTGCGTGGACAGTAAGGATCCGGG

```

### Appendix 3 Sequence alignment of 3'his

Fast alignment of DNA sequences pigh1x.seq and 3'his.seq

Ktuple=2 Gap\_penalty=7

Upper line: pigh1x.seq, from 1 to 1767

Lower line: 3'his.seq, from 10 to 1776

pigh1x.seq:3'his.seq identity= 99.89%(1768/1770) gap=0.51%(9/1779)

```

1 .....ATGAAGTGTCTGCAACTGCTTCTGGCTGGCGTCTCGGCCGTGCTCGCGACT
      |||||||
1 TTTCCATGGATGAAGTGTCTGCAACTGCTTCTGGCTGGCGTCTCGGCCGTGCTCGCGACT

52 GTGAGTGCAGCGCTGTCGTCACCACCTCCGCGAGTGCCAAGCGCTGCTTCCCGGACGAC
   |||||||
61 GTGAGTGCAGCGCTGTCGTCACCACCTCCGCGAGTGCCAAGCGCTGCTTCCCGGACGAC

112 TTCCTCTTTGGCTCTGCCACCGCCGCATACCAAGTGGAGGGCGCTTGGAACGAGACCGGC
   |||||||
121 TTCCTCTTTGGCTCTGCCACCGCCGCATACCAAGTGGAGGGCGCTTGGAACGAGACCGGC

172 CGAACGCCGTCAATCTGGGACGACTTCTGTGCGGAGAAGCCCGGTCTTCAGTGCTCGAAC
   |||||||
181 CGAACGCCGTCAATCTGGGACGACTTCTGTGCGGAGAAGCCCGGTCTTCAGTGCTCGAAC

232 GTCGCCGACGACTTCTACCACCGGTACGTGAGCGACATCCAGACGATGGTCGACTCGGGA
   |||||||
241 GTCGCCGACGACTTCTACCACCGGTACGTGAGCGACATCCAGACGATGGTCGACTCGGGA

292 TTGGACTIONTCCGCTTCAGCATCTCCTGGTCGCGCGCCATGAACTGGGACCCTGCGACC
   |||||||
301 TTGGACTIONTCCGCTTCAGCATCTCCTGGTCGCGCGCCATGAACTGGGACCCTGCGACC

352 AAGAAGATGAAGCCCAACCCGCAAGGCATCGCCTTCTACCACGCGCTCATCGACGAGCTC

```

```

|||||
361  AAGAAGATGAAGCCCAACCCGCAAGGCATCGCCTTCTACCACGCGCTCATCGACGAGCTC
|||||
412  AAGCGCAACAATATCGTGCCGATCTTGACGCTCTACCACTGGGACCTCCCGAGCGAATTG
|||||
421  AAGCGCAACAATATCGTGCCGATCTTGACGCTCTACCACTGGGACCTCCCGAGCGAATTG
|||||
472  CACAGTGAGCTCAAGCCGCAGGGTTGGCTCAACCGCGAGATCATCGATCACTTCGTTGAG
|||||
481  CACAGTGAGCTCAAGCCGCAGGGTTGGCTCAACCGCGAGATCATCGATCACTTCGTTGAG
|||||
532  TACTCGGAGCTCATATTCAACGAGTTTGGCAAGAAGGTGGACATGTGGACGACGTTCAAT
|||||
541  TACTCGGAGCTCATATTCAACGAGTTTGGCAAGAAGGTGGACATGTGGACGACGTTCAAT
|||||
592  GAGCCGCTGTCGTTTCGTCTCCAGGGCTACGCAACGGGCATGTCGGCTCCAGGATACACC
|||||
601  GAGCCGCTGTCGTTTCGTCTCCAGGGCTACGCAACGGGCATGTCGGCTCCAGGATACACC
|||||
652  AACTCCAAGACGAACGCGTACGTTGTGCGGCACAACGTGCTGCTCTCGCACGGCAAGGCG
|||||
661  AACTCCAAGACGAACGCGTACGTTGTGCGGCACAACGTGCTGCTCTCGCACGGCAAGGCG
|||||
712  GTGCAGCGCTTCCGCGAGCTCAAGAAGAGCAACGTCGTCCATGAGAAGGCGCGCATCGGT
|||||
721  GTGCAGCGCTTCCGCGAGCTCAAGAAGAGCAACGTCGTCCATGAGAAGGCGCGCATCGGT
|||||
772  ATCGTGATTGTGCGACCACTTCTACTACCCTCTCGACCCGAGCAACCCGCGCGACGTCATT
|||||
781  ATCGTGATTGTGCGACCACTTCTACTACCCTCTCGACCCGAGCAACCCGCGCGACGTCATT
|||||
832  GCGGCTGAGCGCGCGATCCAGTTTGGATTTCGGCTGGTTCTTCATTCCCATAGTGACTGGC
|||||
841  GCGGCTGAGCGCGCGATCCAGTTTGGATTTCGGCTGGTTCTTCATTCCCATAGTGACTGGC
|||||
892  AAGTATCCAGACGTGATGCGCGAACGTGTGCGGTGACCGCCTGCCAGAGTTCACGGCGGAG
|||||
901  AAGTATCCAGACGTGATGCGCGAACGTGTGCGGTGACCGCCTGCCAGAGTTCACGGCGGAG
|||||
952  GAAGCTGCGATGGTAAAGGACTCGTACGACCTCTTCATGCTCAACCACTACTCATCCAC
|||||
961  GAAGCTGCGATGGTAAAGGACTCGTACGACCTCTTCATGCTCAACCACTACTCATCCAC
|||||
1012 ACTGTGACGGACTGCGACTCCGAAACGTGCTCGTCAAGTGCAGCGATTTGTCCCTCGGA
|||||
1021 ACTGTGACGGACTGCGACTCCGAAACGTGCTCGTCAAGTGCAGCGATTTGTCCCTCGGA
|||||
1072 TGGGAACAGGATATGGGCGTCGACAGTACGCGTGCGCCAGACGGTGCGCGTCTGTCCAGC
|||||
1081 TGGGAACAGGATATGGGCGTCGACAGTACGCGTGCGCCAGACGGTGCGCGTCTGTCCAGC
|||||
1132 AAGAACAGCCTCGGCGAGTACAACGTGCAAGTGGTTACAGGGATACCCGGCAGGCTACCTG
|||||
1141 AAGAACAGCCTCGGCGAGTACAACGTGCAAGTGGTTACAGGGATACCCGGCAGGCTACCTG
|||||
1192 GAAATGATCAAGTGGATGCACAAGTACAACACAAAGGCGGACATCTTACTACCGAGAAC
|||||

```

```

1201 GAAATGATCAAGTGGATGCACAAGTACAACACAAAGGCGGACATCTTACTCACCAGAGAAC
1252 GGCTGGTGC GGCAACGAGACCATCGACAACCAGGACCAACTCTGGTACTTCCAGAGCTAC
      |||
1261 GGCTGGTGC GGCAACGAGACCATCGACAACCAGGACCAACTCTGGTACTTCCAGAGCTAC

1312 ATCGGCGAAGTCTACAAGGCGATCAACGAACCTCAAGATCCCCATCATTGGGTACACTGCG
      |||
1321 ATCGGCGAAGTCTACAAGGCGATCAACGAACCTCAAGATCCCCATCATTGGGTACACTGCG

1372 TGGTCGTTTCATGGACAACCTACGAGTGGGGCTCATTCAAGCCTCGCTTTGGACTCTACTAC
      |||
1381 TGGTCGTTTCATGGACAACCTACGAGTGGGGCTCATTCAAGCCTCGCTTTGGACTCTACTAC

1432 GTCAATTTTCACGTCCGAGACCGGCAAGGCGGATGAGGTACCCCCAGGATCAAACGCCTTG
      |||
1441 GTCAATTTTCACGTCCGAGACCGGCAAGGCGGATGAGGTACCCCCAGGATCAAACGCCTTG

1492 GCACGCATCCCACGTCTCTGCGGCCAAATGGTACGCGCAAGTGGCGAAGACCAAGTGCCTC
      |||
1501 GCACGCATCCCACGTCTCTGCGGCCAAATGGTACGCGCAAGTGGCGAAGACCAAGTGCCTC

1552 GACGGCTTTGAGATTGAAGGCGTGCCAGTCTTGGCAGGAGTCAGCGACGTTACTGGTACC
      |||
1561 GACGGCTTTGAGATTGAAGGCGTGCCAGTCTTGGCAGGAGTCAGCGACGTTACTGGTACC

1612 AGCTCTGGCAGCAGCTTTAAGACGTTTCGCACTCTGGGGGCTGGTGCTGGGTGTTGTTGTC
      |||
1621 AGCTCTGGCAGCAGCTTTAAGACGTTTCGCACTCTGGGGGCTGGTGCTGGGTGTTGTTGTC

1672 GGC GTAC CAGCGATCGCTTTCTTTGTCAAGAGCCGTCGCTCTCGTCAGAGCTCGCTGCTC
      |||
1681 GGC GTAC CAGCGATCGCTTTCTTTGTCAAGAGCCGTCGCTCTCGTCAGAGCTCGCTGCTC

1732 CGCGCGGGCGGAGAACACGCCGCTCGTGCGTGGACAGTAA
      |||

1741 CGCGCGGGCGGAGAACACGCCGCTCGTGCGTGGACAGGAGCAGCCATCATCATCATCATC

```

#### Appendix 4 Sequence alignment of 5'his

Fast alignment of DNA sequences pigh1x.seq and 5'his.seq

Ktuple=2 Gap\_penalty=7

Upper line: pigh1x.seq, from 1 to 1769

Lower line: 5'his.seq, from 67 to 1835

pigh1x.seq:5'his.seq identity= 100.00%(1770/1770) gap=0.00%(0/1770)

```

1   ATGAAGTGTCTGCAACTGCTTCTGGCTGGCGTCTCGGCCGTGCTCGCGACTGTGAGTGCG
      |||
67  ATGAAGTGTCTGCAACTGCTTCTGGCTGGCGTCTCGGCCGTGCTCGCGACTGTGAGTGCG

61  AGCGCTGTGCTCACCACCTCCGCGAGTGCCAAGCGCTGCTTCCCGGACGACTTCCTCTTT

```

```

|||||
127 AGCGCTGTCGTCACACCTCCGCGAGTGCCAAGCGCTGCTTCCCGGACGACTTCCTCTTT
|||||
121 GGCTCTGCCACCGCCGCATACCAAGTGGAGGGCGCTTGGAACGAGACCGGCCGAACGCCG
|||||
187 GGCTCTGCCACCGCCGCATACCAAGTGGAGGGCGCTTGGAACGAGACCGGCCGAACGCCG
|||||
181 TCAATCTGGGACGACTTCTGTCGCGAGAAGCCCGGTCTTCAGTGCTCGAACGTCGCCGAC
|||||
247 TCAATCTGGGACGACTTCTGTCGCGAGAAGCCCGGTCTTCAGTGCTCGAACGTCGCCGAC
|||||
241 GACTTCTACCACCGGTACGTGAGCGACATCCAGACGATGGTCGACTCGGGATTGGACTCG
|||||
307 GACTTCTACCACCGGTACGTGAGCGACATCCAGACGATGGTCGACTCGGGATTGGACTCG
|||||
301 TTCCGCTTCAGCATCTCCTGGTCGCGGCCATGAACTGGGACCCTGCGACCAAGAAGATG
|||||
367 TTCCGCTTCAGCATCTCCTGGTCGCGGCCATGAACTGGGACCCTGCGACCAAGAAGATG
|||||
361 AAGCCCAACCCGCAAGGCATCGCCTTCTACCACGCGCTCATCGACGAGCTCAAGCGCAAC
|||||
427 AAGCCCAACCCGCAAGGCATCGCCTTCTACCACGCGCTCATCGACGAGCTCAAGCGCAAC
|||||
421 AATATCGTGCCGATCTTGACGCTCTACCACTGGGACCTCCCGAGCGAATTGCACAGTGAG
|||||
487 AATATCGTGCCGATCTTGACGCTCTACCACTGGGACCTCCCGAGCGAATTGCACAGTGAG
|||||
481 CTCAAGCCGCAGGGTTGGCTCAACCGCGAGATCATCGATCACTTCGTTGAGTACTCGGAG
|||||
547 CTCAAGCCGCAGGGTTGGCTCAACCGCGAGATCATCGATCACTTCGTTGAGTACTCGGAG
|||||
541 CTCATATTCAACGAGTTTGGCAAGAAGGTGGACATGTGGACGACGTTCAATGAGCCGCTG
|||||
607 CTCATATTCAACGAGTTTGGCAAGAAGGTGGACATGTGGACGACGTTCAATGAGCCGCTG
|||||
601 TCGTTCGTCTCCCAGGGCTACGCAACGGGCATGTGCGCTCCAGGATACACCAACTCCAAG
|||||
667 TCGTTCGTCTCCCAGGGCTACGCAACGGGCATGTGCGCTCCAGGATACACCAACTCCAAG
|||||
661 ACGAACGCGTACGTTGTGCGGCACAACGTGCTGCTCTCGCACGGCAAGGCGGTGCAGCGC
|||||
727 ACGAACGCGTACGTTGTGCGGCACAACGTGCTGCTCTCGCACGGCAAGGCGGTGCAGCGC
|||||
721 TTCCGCGAGCTCAAGAAGAGCAACGTGCTCCATGAGAAGGCGCGCATCGGTATCGTGATT
|||||
787 TTCCGCGAGCTCAAGAAGAGCAACGTGCTCCATGAGAAGGCGCGCATCGGTATCGTGATT
|||||
781 GTCGACCACTTCTACTACCCTCTCGACCCGAGCAACCCGCGCGACGTCATTGCGGCTGAG
|||||
847 GTCGACCACTTCTACTACCCTCTCGACCCGAGCAACCCGCGCGACGTCATTGCGGCTGAG
|||||
841 CGCGCGATCCAGTTTGGATTTCGGCTGGTTCTTCATTCCCATAGTGACTGGCAAGTATCCA
|||||
907 CGCGCGATCCAGTTTGGATTTCGGCTGGTTCTTCATTCCCATAGTGACTGGCAAGTATCCA
|||||
901 GACGTGATGCGCGAACGTGTGCGGTGACCGCCTGCCAGAGTTCACGGCGGAGGAAGCTGCG
|||||

```

967 GACGTGATGCGCGAACGTGTCGGTGACCGCCTGCCAGAGTTCACGGCGGAGGAAGCTGCG  
 961 ATGGTAAAGGACTCGTACGACCTCTTCATGCTCAACCACTACTCATCCCACACTGTGACG  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1027 ATGGTAAAGGACTCGTACGACCTCTTCATGCTCAACCACTACTCATCCCACACTGTGACG  
 1021 GACTGCGACTCCGAAACGTGCTCGTCAAGTGCAGCGATTTGTCCCTCGGATGGGAACAG  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1087 GACTGCGACTCCGAAACGTGCTCGTCAAGTGCAGCGATTTGTCCCTCGGATGGGAACAG  
 1081 GATATGGGCGTCGACAGTACGCGTGCGCCAGACGGTGC GCGTCTGTCCAGCAAGAACAGC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1147 GATATGGGCGTCGACAGTACGCGTGCGCCAGACGGTGC GCGTCTGTCCAGCAAGAACAGC  
 1141 CTCGGCGAGTACAAGTGAAGTGGTTACGGGATACCCGGCAGGCTACCTGGAAATGATC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1207 CTCGGCGAGTACAAGTGAAGTGGTTACGGGATACCCGGCAGGCTACCTGGAAATGATC  
 1201 AAGTGGATGCACAAGTACAACACAAAGGCGGACATCTTACTCACCGAGAACGGCTGGTGC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1267 AAGTGGATGCACAAGTACAACACAAAGGCGGACATCTTACTCACCGAGAACGGCTGGTGC  
 1261 GGCAACGAGACCATCGACAACCAGGACCAACTCTGGTACTTCCAGAGCTACATCGGCGAA  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1327 GGCAACGAGACCATCGACAACCAGGACCAACTCTGGTACTTCCAGAGCTACATCGGCGAA  
 1321 GTCTACAAGGCGATCAACGAACTCAAGATCCCCATCATTGGGTACACTGCGTGGTCGTTC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1387 GTCTACAAGGCGATCAACGAACTCAAGATCCCCATCATTGGGTACACTGCGTGGTCGTTC  
 1381 ATGGACAACCTACGAGTGGGGCTCATTCAAGCCTCGCTTTGGACTCTACTACGTCAATTTT  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1447 ATGGACAACCTACGAGTGGGGCTCATTCAAGCCTCGCTTTGGACTCTACTACGTCAATTTT  
 1441 ACGTCCGAGACCGGCAAGGCGGATGAGGTACCCCCAGGATCAAACGCCTTGGCAGGCATC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1507 ACGTCCGAGACCGGCAAGGCGGATGAGGTACCCCCAGGATCAAACGCCTTGGCAGGCATC  
 1501 CCACGTCCTGCGGCCAAATGGTACGCGCAAGTGGCGAAGACCAAGTGCCTCGACGGCTTT  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1567 CCACGTCCTGCGGCCAAATGGTACGCGCAAGTGGCGAAGACCAAGTGCCTCGACGGCTTT  
 1561 GAGATTGAAGGCGTGCCAGTCTTGGCAGGAGTCAGCGACGTTACTGGTACCAGCTCTGGC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1627 GAGATTGAAGGCGTGCCAGTCTTGGCAGGAGTCAGCGACGTTACTGGTACCAGCTCTGGC  
 1621 AGCAGCTTTAAGACGTTTCGCACTCTGGGGGCTGGTGCTGGGTGTTGTTGTTCGGCGTACCA  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1687 AGCAGCTTTAAGACGTTTCGCACTCTGGGGGCTGGTGCTGGGTGTTGTTGTTCGGCGTACCA  
 1681 GCGATCGCTTTTCTTTGTCAAGAGCCGTCGCTCTCGTCAGAGCTCGCTGCTCCGCGCGGGC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1747 GCGATCGCTTTTCTTTGTCAAGAGCCGTCGCTCTCGTCAGAGCTCGCTGCTCCGCGCGGGC  
 1741 GAGAACACGCCGCTCGTGCGTGGACAGTAA  
 ||||||||||||||||||||||||||||||||  
 1807 GAGAACACGCCGCTCGTGCGTGGACAGTAAGGATCCGGGA

## Appendix 5 Sequence alignment of *short*

Fast alignment of DNA sequences short.seq and tPiGH1-x.seq

Ktuple=2 Gap\_penalty=7

Upper line: short.seq, from 1 to 1799

Lower line: tPiGH1-x.seq, from 1 to 1799

short.seq:tPiGH1-x.seq identity= 100.00%(1800/1800) gap=0.00%(0/1800)

```

1      TTTTCCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCGAAAAC TTGTATTTCT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1      TTTTCCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCGAAAAC TTGTATTTCT

61     CAAAGTAGCGCTGTCGTCACCACCTCCGCGAGTGCCAAGCGCTGCTTCCCGGACGACTTC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61     CAAAGTAGCGCTGTCGTCACCACCTCCGCGAGTGCCAAGCGCTGCTTCCCGGACGACTTC

121    CTCTTTGGCTCTGCCACCGCCGCATACCAAGTGGAGGGCGCTTGGAACGAGACCGGCCGA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121    CTCTTTGGCTCTGCCACCGCCGCATACCAAGTGGAGGGCGCTTGGAACGAGACCGGCCGA

181    ACGCCGTCAATCTGGGACGACTTCTGTGTCGCGAGAAGCCCGGTCTTCAGTGCTCGAACGTC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181    ACGCCGTCAATCTGGGACGACTTCTGTGTCGCGAGAAGCCCGGTCTTCAGTGCTCGAACGTC

241    GCCGACGACTTCTACCACCGGTACGTGAGCGACATCCAGACGATGGTCGACTCGGGATTG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241    GCCGACGACTTCTACCACCGGTACGTGAGCGACATCCAGACGATGGTCGACTCGGGATTG

301    GACTCGTTCCGCTTCAGCATCTCCTGGTTCGCGCGCCATGAACTGGGACCCTGCGACCAAG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301    GACTCGTTCCGCTTCAGCATCTCCTGGTTCGCGCGCCATGAACTGGGACCCTGCGACCAAG

361    AAGATGAAGCCCAACCCGCAAGGCATCGCCTTCTACCACGCGCTCATCGACGAGCTCAAG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361    AAGATGAAGCCCAACCCGCAAGGCATCGCCTTCTACCACGCGCTCATCGACGAGCTCAAG

421    CGCAACAATATCGTGCCGATCTTGACGCTCTACCACTGGGACCTCCCGAGCGAATTGCAC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421    CGCAACAATATCGTGCCGATCTTGACGCTCTACCACTGGGACCTCCCGAGCGAATTGCAC

481    AGTGAGCTCAAGCCGCAGGGTTGGCTCAACCGCGAGATCATCGATCACTTCGTTGAGTAC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481    AGTGAGCTCAAGCCGCAGGGTTGGCTCAACCGCGAGATCATCGATCACTTCGTTGAGTAC

541    TCGGAGCTCATATTCAACGAGTTTGGCAAGAAGGTGGACATGTGGACGACGTTCAATGAG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541    TCGGAGCTCATATTCAACGAGTTTGGCAAGAAGGTGGACATGTGGACGACGTTCAATGAG

601    CCGCTGTCTGTTCTGCTCTCCAGGGCTACGCAACGGGCATGTCTGGCTCCAGGATACACCAAC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601    CCGCTGTCTGTTCTGCTCTCCAGGGCTACGCAACGGGCATGTCTGGCTCCAGGATACACCAAC

```

```

661  TCCAAGACGAACGCGTACGTTGTCGCGCACAAACGTGCTGCTCTCGCACGGCAAGGCGGTG
    |
661  TCCAAGACGAACGCGTACGTTGTCGCGCACAAACGTGCTGCTCTCGCACGGCAAGGCGGTG

721  CAGCGCTTCCGCGAGCTCAAGAAGAGCAACGTGCTCCATGAGAAGGCGCGCATCGGTATC
    |
721  CAGCGCTTCCGCGAGCTCAAGAAGAGCAACGTGCTCCATGAGAAGGCGCGCATCGGTATC

781  GTGATTGTCGACCACTTCTACTACCCTCTCGACCCGAGCAACCCGCGCGACGTCATTGCG
    |
781  GTGATTGTCGACCACTTCTACTACCCTCTCGACCCGAGCAACCCGCGCGACGTCATTGCG

841  GCTGAGCGCGCGATCCAGTTTGGATTTCGGCTGGTTCTTCATTCCCATAGTGACTGGCAAG
    |
841  GCTGAGCGCGCGATCCAGTTTGGATTTCGGCTGGTTCTTCATTCCCATAGTGACTGGCAAG

901  TATCCAGACGTGATGCGCGAACGTGTCGGTGACCGCCTGCCAGAGTTCACGGCGGAGGAA
    |
901  TATCCAGACGTGATGCGCGAACGTGTCGGTGACCGCCTGCCAGAGTTCACGGCGGAGGAA

961  GCTGCGATGGTAAAGGACTCGTACGACCTCTTCATGCTCAACCACTACTCATCCCACACT
    |
961  GCTGCGATGGTAAAGGACTCGTACGACCTCTTCATGCTCAACCACTACTCATCCCACACT

1021 GTGACGGACTGCGACTCCGAAACGTGCTCGTCAAGTGCAGCGATTTGTCCCTCGGATGG
    |
1021 GTGACGGACTGCGACTCCGAAACGTGCTCGTCAAGTGCAGCGATTTGTCCCTCGGATGG

1081 GAACAGGATATGGGCGTCGACAGTACGCGTGCGCCAGACGGTGCGCGTCTGTCCAGCAAG
    |
1081 GAACAGGATATGGGCGTCGACAGTACGCGTGCGCCAGACGGTGCGCGTCTGTCCAGCAAG

1141 AACAGCCTCGGCGAGTACAAGTGAAGTGGTTCACGGGATACCCGGCAGGCTACCTGGAA
    |
1141 AACAGCCTCGGCGAGTACAAGTGAAGTGGTTCACGGGATACCCGGCAGGCTACCTGGAA

1201 ATGATCAAGTGGATGCACAAGTACAACACAAAGGCGGACATCTTACTCACCGAGAACGGC
    |
1201 ATGATCAAGTGGATGCACAAGTACAACACAAAGGCGGACATCTTACTCACCGAGAACGGC

1261 TGGTGCGGCAACGAGACCATCGACAACCAGGACCAACTCTGGTACTTCCAGAGCTACATC
    |
1261 TGGTGCGGCAACGAGACCATCGACAACCAGGACCAACTCTGGTACTTCCAGAGCTACATC

1321 GGCGAAGTCTACAAGGCGATCAACGAACTCAAGATCCCCATCATTGGGTACACTGCGTGG
    |
1321 GGCGAAGTCTACAAGGCGATCAACGAACTCAAGATCCCCATCATTGGGTACACTGCGTGG

1381 TCGTTTCATGGACAACTACGAGTGGGGCTCATTCAAGCCTCGCTTTGGACTCTACTACGTC
    |
1381 TCGTTTCATGGACAACTACGAGTGGGGCTCATTCAAGCCTCGCTTTGGACTCTACTACGTC

1441 AATTTTCAGTCCGAGACCGGCAAGGCGGATGAGGTCACCCCAGGATCAAACGCCTTGGCA
    |
1441 AATTTTCAGTCCGAGACCGGCAAGGCGGATGAGGTCACCCCAGGATCAAACGCCTTGGCA

1501 CGCATCCCACGTCCTGCGGCCAAATGGTACGCGCAAGTGGCGAAGACCAAGTGCCTCGAC

```

```
|||||
1501 CGCATCCCACGTCCTGCGGCCAAATGGTACGCGCAAGTGGCGAAGACCAAGTGCCTCGAC
|||||
1561 GGCTTTGAGATTGAAGGCGTGCCAGTCTTGGCAGGAGTCAGCGACGTTACTGGTACCAGC
|||||
1561 GGCTTTGAGATTGAAGGCGTGCCAGTCTTGGCAGGAGTCAGCGACGTTACTGGTACCAGC
|||||
1621 TCTGGCAGCAGCTTTAAGACGTTGCACTCTGGGGGCTGGTGCTGGGTGTTGTTGTCGGC
|||||
1621 TCTGGCAGCAGCTTTAAGACGTTGCACTCTGGGGGCTGGTGCTGGGTGTTGTTGTCGGC
|||||
1681 GTACCAGCGATCGCTTTCTTTGTCAAGAGCCGTCGCTCTCGTCAGAGCTCGCTGCTCCGC
|||||
1681 GTACCAGCGATCGCTTTCTTTGTCAAGAGCCGTCGCTCTCGTCAGAGCTCGCTGCTCCGC
|||||
1741 GCGGGCGAGAACACGCCGCTCGTGCGTGGACAGTAAGGATCCGGA
|||||
1741 GCGGGCGAGAACACGCCGCTCGTGCGTGGACAGTAAGGATCCGGA
```



## 8 CURRICULUM VITAE

### **Post-secondary Education**

Graduate Degree MSc in Biology (2015 - Present)

University of Western Ontario, London, ON, Canada

Undergraduate Degree B.Sc. in Biotechnology Engineering (2009 - 2013)

Instituto Tecnológico y de Estudios Superiores de Monterrey, Monterrey, N.L., Mexico

### **Professional Experience**

Teaching Assistant (2015 - 2017)

University of Western Ontario – Department of Biology

Research Assistant (2015)

Instituto Tecnológico y de Estudios Superiores de Monterrey- MolBree Research Group,  
Monterrey, N.L., Mexico

Health Inspector (2014-2015)

Health Secretariat of the State of Nuevo Leon, Monterrey, N.L., Mexico

Research Assistant (2012-2013)

Instituto Tecnológico y de Estudios Superiores de Monterrey- Nutraceuticals Research  
Group, Monterrey, N.L., Mexico

## List of Publications and Presentations

### Non-refereed contributions (Poster presentations)

**Puebla, S.\*** and Bernards, M.A. (2016) Functional and biochemical characterization of a putative ginsenosidase from the oomycete *Pythium irregulare*. Presented at the Symposium on Sustainable Agriculture, University of Western Ontario, April 14 \*Poster Presenter.

**Puebla, S.\*** and Bernards, M.A. (2016) Functional and biochemical characterization of a putative ginsenosidase from the oomycete *Pythium irregulare*. Presented at the Biology Graduate Research Forum, University of Western Ontario, October 13 and 14 \*Poster Presenter.

**Puebla, S.\*** and Garcia-Lara, S. (2014) Evaluation of the edibility and nutritional value of the insect *Sitophilus zeamais*. Presented at the 1<sup>st</sup> International Congress of Functional Foods and Nutraceuticals, Instituto Tecnológico y de Estudios Superiores de Monterrey, June 11-13 \*Poster Presenter.

## Scholarships

CONACyT Scholarship (2015-2017)